CHAPTER 3

INTRODUCTION TO TRANSMISSION LINES AND WAVEGUIDES

A TRANSMISSION LINE is a device designed to guide electrical energy from one point to another. It is used, for example, to transfer the output rf energy of a transmitter to an antenna. This energy will not travel through normal electrical wire without great losses. Although the antenna can be connected directly to the transmitter, the antenna is usually located some distance away from the transmitter. On board ship, the transmitter is located inside a radio room, and its associated antenna is mounted on a mast. A transmission line is used to connect the transmitter and the antenna.

The transmission line has a single purpose for both the transmitter and the antenna. This purpose is to transfer the energy output of the transmitter to the antenna with the least possible power loss. How well this is done depends on the special physical and electrical characteristics (impedance and resistance) of the transmission line.

TRANSMISSION LINE THEORY

The electrical characteristics of a two-wire transmission line depend primarily on the construction of the line. The two-wire line acts like a long capacitor. The change of its capacitive reactance is noticeable as the frequency applied to it is changed. Since the long conductors have a magnetic field about them when electrical energy is being passed through them, they also exhibit the properties of inductance. The values of inductance and capacitance presented depend on the various physical factors that we discussed earlier. For example, the type of line used, the dielectric in the line, and the length of the line must be considered. The effects of the inductive and capacitive reactance of the line depend on the frequency applied. Since no dielectric is perfect, electrons manage to move from one conductor to the other through the dielectric. Each type of two-wire transmission line also has a conductance value. This conductance value represents the value of the current flow that may be expected through the insulation, If the line is uniform (all values equal at each unit length), then one small section of the line may represent several feet. This illustration of a two-wire transmission line will be used throughout the discussion of transmission lines; but, keep in mind that the principles presented apply to all transmission lines. We will explain the theories using LUMPED CONSTANTS and DISTRIBUTED CONSTANTS to further simplify these principles.

LUMPED CONSTANTS

A transmission line has the properties of inductance, capacitance, and resistance just as the more conventional circuits have. Usually, however, the constants in conventional circuits are lumped into a single device or component. For example, a coil of wire has the property of inductance. When a certain amount of inductance is needed in a circuit, a coil of the proper dimensions is inserted. The inductance of the circuit is lumped into the one component. Two metal plates separated by a small space, can be used to supply the required capacitance for a circuit. In such a case, most of the capacitance of the circuit is lumped into this one component. Similarly, a fixed resistor can be used to supply a certain value of circuit resistance as a lumped sum. Ideally, a transmission line would also have its constants of inductance, capacitance, and resistance lumped together, as shown in figure 3-1. Unfortunately, this is not the case. Transmission line constants are as described in the following paragraphs.

DISTRIBUTED CONSTANTS

Transmission line constants, called distributed constants, are spread along the entire length of the transmission line and cannot be distinguished separately. The amount of inductance, capacitance, and resistance depends on the length of the line, the size of the conducting wires, the spacing between the

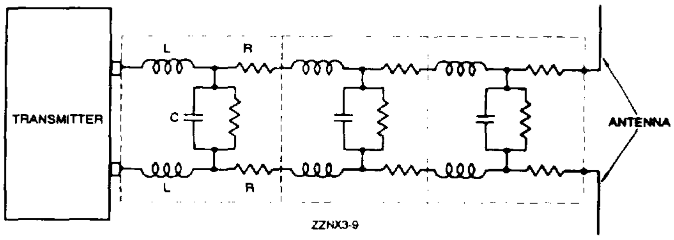

Figure 3-1.—Two-wire transmission line.

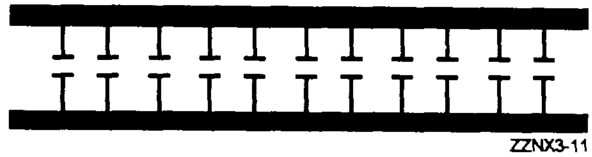

Figure 3-3.—Distributed capacitance.

wires, and the dielectric (air or insulating medium) between the wires. The following paragraphs will be useful to you as you study distributed constants on a transmission line.

Inductance of a Transmission Line

When current flows through a wire, magnetic lines of force are set up around the wire. As the current increases and decreases in amplitude, the field around the wire expands and collapses accordingly. The energy produced by the magnetic lines of force collapsing back into the wire tends to keep the current flowing in the same direction. This represents a certain amount of inductance, which is expressed in microhenrys per unit length. Figure 3-2 illustrates the inductance and magnetic fields of a transmission line.

Capacitance of a Transmission Line

Capacitance also exists between the transmission line wires, as illustrated in figure 3-3. Notice that the two parallel wires act as plates of a capacitor and that the air between them acts as a dielectric. The capacitance between the wires is usually expressed in picofarads per unit length. This electric field between the wires is similar to the field that exists between the two plates of a capacitor.

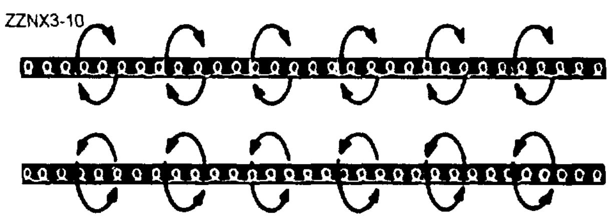

Figure 3-2.—Distributed inductance.

Resistance of a Transmission Line

The transmission line shown in figure 3-4 has electrical resistance along its length. This resistance is usually expressed in ohms per unit length and is shown as existing continuously from one end of the line to the other.

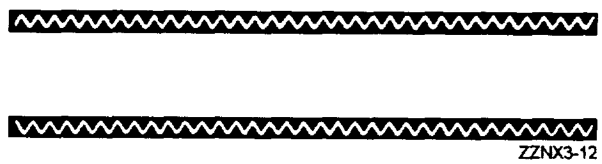

Figure 3-4.—Distributed resistance.

Leakage Current

Since any dielectric, even air, is not a perfect insulator, a small current known as LEAKAGE CURRENT flows between the two wires. In effect, the insulator acts as a resistor, permitting current to pass between the two wires. Figure 3-5 shows this leakage path as resistors in parallel connected between the two lines. This property is called CONDUCTANCE (G) and is the opposite of resistance. Conductance in transmission lines is expressed as the reciprocal of resistance and is usually given in micromhos per unit length.

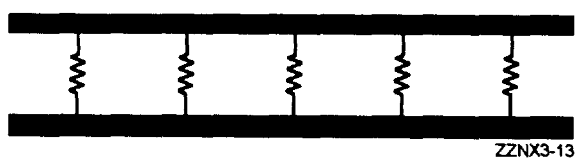

Figure 3-5.—Leakage in a transmission line.

ELECTROMAGNETIC FIELDS

The distributed constants of resistance, inductance, and capacitance are basic properties common to all transmission lines and exist whether or not any current flow exists. As soon as current flow and voltage exist in a transmission line, another property becomes quite evident. This is the presence of an electromagnetic field, or lines of force, about the wires of the transmission line. The lines of force themselves are not visible; however, understanding the force that an electron experiences while in the field of these lines is very important to your understanding of energy transmission.

There are two kinds of fields; one is associated with voltage and the other with current. The field associated with voltage is called the ELECTRIC (E) FIELD. It exerts a force on any electric charge placed in it. The field associated with current is called a MAGNETIC (H) FIELD, because it tends to exert a force on any magnetic pole placed in it. Figure 3-6 illustrates the way in which the E fields and H fields tend to orient themselves between conductors of a typical two-wire transmission line. The illustration shows a cross section of the transmission lines. The E field is represented by solid lines and the H field by dotted lines. The arrows indicate the direction of the lines of force. Both fields normally exist together and are spoken of collectively as the electromagnetic field.

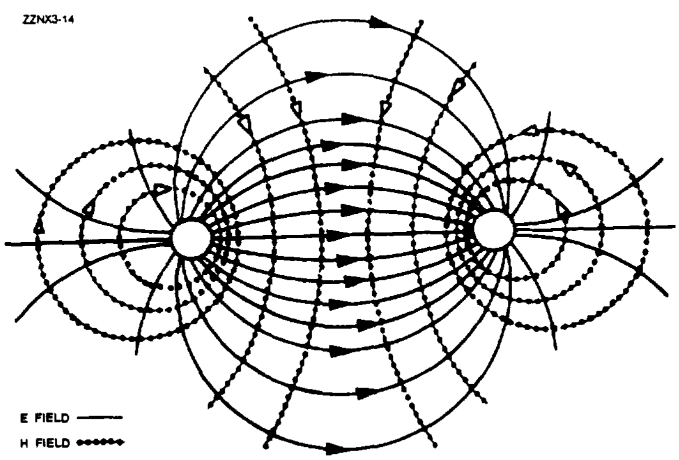

Figure 3-6.—Fields between conductors.

CHARACTERISTIC IMPEDANCE

You can describe a transmission line in terms of its impedance. The ratio of voltage to current (E_{in}/I_{in}) at the input end is known as the INPUT IMPEDANCE (Z_{in}). This is the impedance presented to the transmitter by the transmission line and its load, the antenna. The ratio of voltage to current at the output (E_{OUT}/I_{OUT}) end is known as the OUTPUT IMPEDANCE (Z_{OUT}). This is the impedance presented to the load by the transmission line and its source. If an infinitely long transmission line could be used, the ratio of voltage to current at any point on that transmission line would be some particular value of impedance. This impedance is known as the CHARACTERISTIC IMPEDANCE.

The maximum (and most efficient) transfer of electrical energy takes place when the source impedance is matched to the load impedance. This fact is very important in the study of transmission lines and antennas. If the characteristic impedance of the transmission line and the load impedance are equal, energy from the transmitter will travel down the transmission line to the antenna with no power loss caused by reflection.

LINE LOSSES

The discussion of transmission lines so far has not directly addressed LINE LOSSES; actually some losses occur in all lines. Line losses may be any of three types—COPPER, DIELECTRIC, and RADIATION or INDUCTION LOSSES.

NOTE: Transmission lines are sometimes referred to as rf lines. In this text the terms are used interchangeably.

Copper Losses

One type of copper loss is I^2R LOSS. In rf lines the resistance of the conductors is never equal to zero. Whenever current flows through one of these conductors, some energy is dissipated in the form of heat. This heat loss is a POWER LOSS. With copper braid, which has a resistance higher than solid tubing, this power loss is higher.

Another type of copper loss is due to SKIN EFFECT. When dc flows through a conductor, the movement of electrons through the conductor's cross section is uniform, The situation is somewhat different when ac is applied. The expanding and collapsing fields about each electron encircle other electrons. This phenomenon, called SELF INDUCTION, retards the movement of the encircled electrons. The flux density at the center is so great that electron movement at this point is reduced. As frequency is increased, the opposition to the flow of current in the center of the wire increases. Current in the center of the wire becomes smaller and most of the electron flow is on the wire surface. When the frequency applied is 100 megahertz or higher, the electron movement in the center is so small that the center of the wire could be removed without any noticeable effect on current. You should be able to see that the effective cross-sectional area decreases as the frequency increases. Since resistance is inversely proportional to the cross-sectional area, the resistance will increase as the frequency is increased. Also, since power loss increases as resistance increases, power losses increase with an increase in frequency because of skin effect.

Copper losses can be minimized and conductivity increased in an rf line by plating the line with silver. Since silver is a better conductor than copper, most of the current will flow through the silver layer. The tubing then serves primarily as a mechanical support.

Dielectric Losses

DIELECTRIC LOSSES result from the heating effect on the dielectric material between the conductors. Power from the source is used in heating the dielectric. The heat produced is dissipated into the surrounding medium. When there is no potential difference between two conductors, the atoms in the dielectric material between them are normal and the orbits of the electrons are circular. When there is a potential difference between two conductors, the orbits of the electrons change. The excessive negative charge on one conductor repels electrons on the dielectric toward the positive conductor and thus distorts the orbits of the electrons. A change in the path of electrons requires more energy, introducing a power loss.

The atomic structure of rubber is more difficult to distort than the structure of some other dielectric materials. The atoms of materials, such as polyethylene, distort easily. Therefore, polyethylene is often used as a dielectric because less power is consumed when its electron orbits are distorted.

Radiation and Induction Losses

RADIAION and INDUCTION LOSSES are similar in that both are caused by the fields surrounding the conductors. Induction losses occur when the electromagnetic field about a conductor cuts through any nearby metallic object and a current is induced in that object. As a result, power is dissipated in the object and is lost.

Radiation losses occur because some magnetic lines of force about a conductor do not return to the conductor when the cycle alternates. These lines of force are projected into space as radiation, and this results in power losses. That is, power is supplied by the source, but is not available to the load.

VOLTAGE CHANGE

In an electric circuit, energy is stored in electric and magnetic fields. These fields must be brought to the load to transmit that energy. At the load, energy contained in the fields is converted to the desired form of energy.

Transmission of Energy

When the load is connected directly to the source of energy, or when the transmission line is short, problems concerning current and voltage can be solved by applying Ohm's law. When the transmission line becomes long enough so the time difference between a change occurring at the generator and a change appearing at the load becomes appreciable, analysis of the transmission line becomes important.

Dc Applied to a Transmission Line

In figure 3-7, a battery is connected through a relatively long two-wire transmission line to a load at the far end of the line. At the instant the switch is closed, neither current nor voltage exists on the line. When the switch is closed, point A becomes a positive potential, and point B becomes negative. These points of difference in potential move down the line. However, as the initial points of potential leave points A and B, they are followed by new points of difference in potential, which the battery adds at A and B. This is merely saying that the battery maintains a constant potential difference between points A and B. A short time after the switch is closed, the initial points of difference in potential have reached points A' and B'; the wire sections from points A to A' and points B to B' are at the same potential as A and B, respectively. The points of charge are represented by plus (+) and minus (-) signs along the wires, The directions of the currents in the wires are represented by the arrowheads on the line, and the direction of travel is indicated by an arrow below the line. Conventional lines of force represent the electric field that exists between the opposite kinds of charge on the wire sections from A to A' and B to B'. Crosses (tails of arrows) indicate the magnetic field created by the electric field moving down the line. The moving electric field and the accompanying magnetic field constitute an electromagnetic wave that is moving from the generator (battery) toward the load. This wave travels at approximately the speed of light in free space. The energy reaching the load is equal to that developed at the battery (assuming there are no losses in the transmission line). If the load absorbs all of the energy, the current and voltage will be evenly distributed along the line.

Ac Applied to a Transmission Line

When the battery of figure 3-7 is replaced by an ac generator (fig. 3-8), each successive instantaneous value of the generator voltage is propagated down the

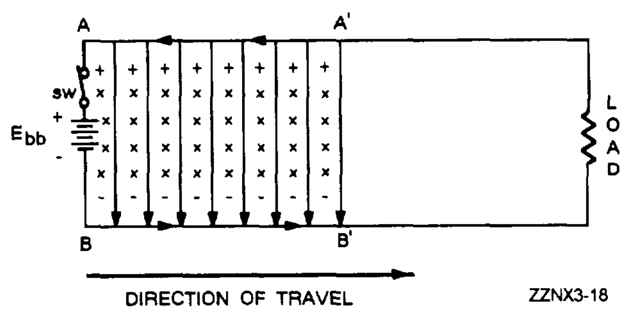

Figure 3-7.—Dc voltage applied to a line.

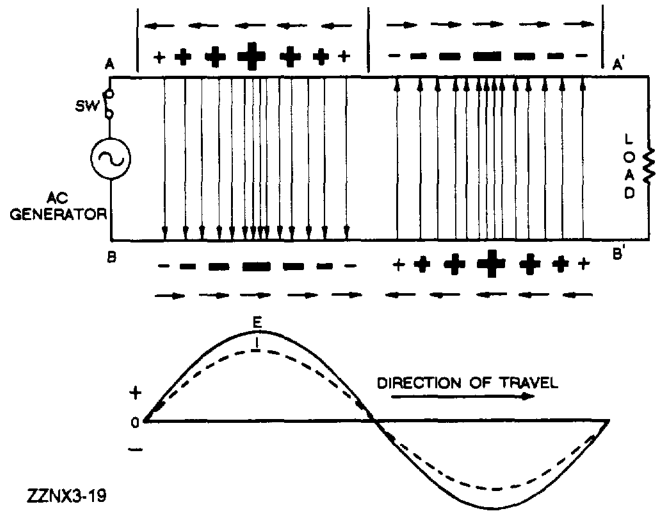

Figure 3-8.—Ac voltage applied to a line.

line at the speed of light. The action is similar to the wave created by the battery, except the applied voltage is sinusoidal instead of constant. Assume that the switch is closed at the moment the generator voltage is passing through zero and that the next half cycle makes point A positive. At the end of one cycle of generator voltage, the current and voltage distribution will be as shown in figure 3-8.

In this illustration the conventional lines of force represent the electric fields. For simplicity, the magnetic fields are not shown. Points of charge are indicated by plus (+) and minus (-) signs, the larger signs indicating points of higher amplitude of both voltage and current. Short arrows indicate direction of current (electron flow). The waveform drawn below the transmission line represents the voltage (E) and current (I) waves. The line is assumed to be infinite in length so there is no reflection. Thus, traveling sinusoidal voltage and current waves continually travel in phase from the generator toward the load, or far end of the line. Waves traveling from the generator to the load are called INCIDENT WAVES. Waves traveling from the load back to the generator are called REFLECTED WAVES and will be explained in later paragraphs.

STANDING-WAVE RATIO

The measurement of standing waves on a transmission line yields information about equipment operating conditions. Maximum power is absorbed by the load when $Z_L = Z_0$. If a line has no standing waves, the termination for that line is correct and maximum power transfer takes place.

You have probably noticed that the variation of standing waves shows how near the rf line is to being terminated in Z_0. A wide variation in voltage along the length means a termination far from Z_0. A small variation means termination near Z_0. Therefore, the ratio of the maximum to the minimum is a measure of the perfection of the termination of a line. This ratio is called the STANDING-WAVE RATIO (swr) and is always expressed in whole numbers. For example, a ratio of 1:1 describes a line terminated in its characteristic impedance (Z_0).

Voltage Standing-Wave Ratio

The ratio of maximum voltage to minimum voltage on a line is called the VOLTAGE STANDING-WAVE RATIO (vswr). Therefore:

$$\text{vswr} = \left| \frac{E_{max}}{E_{min}} \right|$$

The vertical lines in the formula indicate that the enclosed quantities are absolute and that the two values are taken without regard to polarity, Depending on the nature of the standing waves, the numerical value of vswr ranges from a value of 1 ($Z_L = Z_0$, no standing waves) to an infinite value for theoretically complete reflection. Since there is always a small loss on a line, the minimum voltage is never zero and the vswr is always some finite value. However, if the vswr is to be a useful quantity. the power losses along the line must be small in comparison to the transmitted power.

Power Standing-Wave Ratio

The square of the vswr is called the POWER STANDING-WAVE RATIO (pswr). Therefore:

$$\text{pswr} = \frac{P_{max}}{P_{min}}$$

This ratio is useful because the instruments used to detect standing waves react to the square of the voltage. Since power is proportional to the square of the voltage, the ratio of the square of the maximum and minimum voltages is called the power standing-wave ratio. In a sense, the name is misleading because the power along a transmission line does not vary.

Current Standing-Wave Ratio

The ratio of maximum to minimum current along a transmission line is called CURRENT STANDING- WAVE RATIO (iswr). Therefore:

$$\text{iswr} = \left| \frac{I_{max}}{I_{min}} \right|$$

This ratio is the same as that for voltages. It can be used where measurements are made with loops that sample the magnetic field along a line. It gives the same results as vswr measurements.

TRANSMISSION MEDIUMS

The Navy uses many different types of TRANSMISSION MEDIUMS in its electronic applications. Each medium (line or waveguide) has a certain characteristic impedance value, current-carrying capacity, and physical shape and is designed to meet a particular requirement.

The five types of transmission mediums that we will discuss in this topic include PARALLEL-LINE, TWISTED PAIR, SHIELDED PAIR, COAXIAL LINE, and WAVEGUIDES. The use of a particular line depends, among other things, on the applied frequency, the power-handling capabilities, and the type of installation.

Parallel Line

One type of parallel line is the TWO-WIRE OPEN LINE, illustrated in figure 3-9. This line consists of two wires that are generally spaced from 2 to 6 inches apart by insulating spacers. This type of line is most often used for power lines, rural telephone lines, and telegraph lines. It is sometimes used as a transmission line between a transmitter and an antenna or between an antenna and a receiver. An advantage of this type

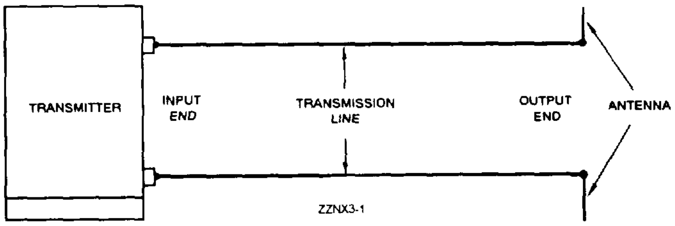

Figure 3-9.—Two-wire open line.

Figure 3-11.—Twisted pair.

of line is its simple construction. The principal disadvantages of this type of line are the high radiation losses and electrical noise pickup because of the lack of shielding. Radiation losses are produced by the changing fields created by the changing current in each conductor.

Another type of parallel line is the TWO-WIRE RIBBON (TWIN LEAD) LINE, illustrated in figure 3-10. This type of transmission line is commonly used to connect a television receiving antenna to a home television set. This line is essentially the same as the two-wire open line except that uniform spacing is assured by embedding the two wires in a low-loss dielectric, usually polyethylene. Since the wires are embedded in the thin ribbon of polyethylene, the dielectric space is partly air and partly polyethylene.

Twisted Pair

The TWISTED PAIR transmission line is illustrated in figure 3-11. As the name implies, the line consists of two insulated wires twisted together to form a flexible line without the use of spacers. It is not used for transmitting high frequency because of the high dielectric losses that occur in the rubber insulation. When the line is wet, the losses increase greatly.

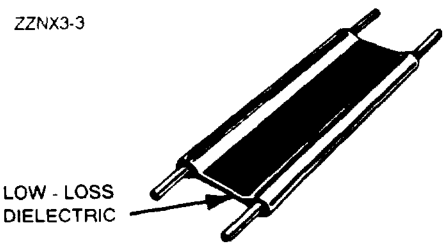

Figure 3-10.—Two-wire ribbon line.

Shielded Pair

The SHIELDED PAIR, shown in figure 3-12, consists of parallel conductors separated from each other and surrounded by a solid dielectric. The conductors are contained within a braided copper tubing that acts as an electrical shield. The assembly is covered with a rubber or flexible composition coating that protects the line from moisture and mechanical damage. Outwardly, it looks much like the power cord of a washing machine or refrigerator.

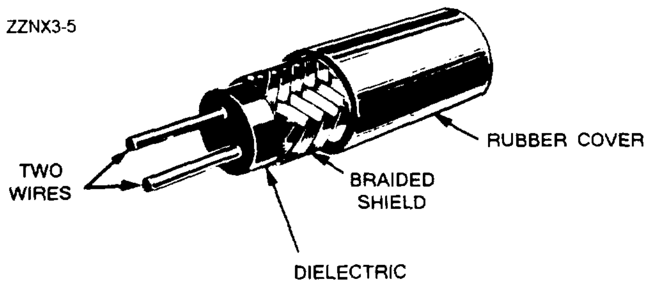

Figure 3-12.—Shielded pair.

The principal advantage of the shielded pair is that the conductors are balanced to ground; that is, the capacitance between the wires is uniform throughout the length of the line. This balance is due to the uniform spacing of the grounded shield that surrounds the wires along their entire length. The braided copper shield isolates the conductors from stray magnetic fields.

Coaxial Lines

There are two types of COAXIAL LINES, RIGID (AIR) COAXIAL LINE and FLEXIBLE (SOLID) COAXIAL LINE. The physical construction of both types is basically the same; that is, each contains two concentric conductors.

The rigid coaxial line consists of a central, insulated wire (inner conductor) mounted inside a tubular outer conductor. This line is shown in figure 3-13. In some applications, the inner conductor is also tubular. The inner conductor is insulated from the outer conductor by insulating spacers or beads at regular intervals. The spacers are made of Pyrex, polystyrene, or some other material that has good insulating characteristics and low dielectric losses at high frequencies.

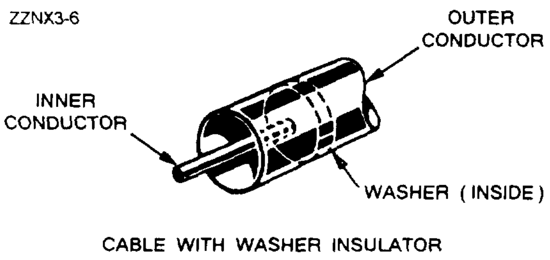

Figure 3-13.—Air coaxial line.

The chief advantage of the rigid line is its ability to minimize radiation losses. The electric and magnetic fields in a two-wire parallel line extend into space for relatively great distances and radiation losses occur. However, in a coaxial line no electric or magnetic fields extend outside of the outer conductor. The fields are confined to the space between the two conductors, resulting in a perfectly shielded coaxial line. Another advantage is that interference from other lines is reduced.

The rigid line has the following disadvantages: (1) it is expensive to construct; (2) it must be kept dry to prevent excessive leakage between the two conductors; and (3) although high-frequency losses are somewhat less than in previously mentioned lines, they are still excessive enough to limit the practical length of the line.

Leakage caused by the condensation of moisture is prevented in some rigid line applications by the use of an inert gas, such as nitrogen, helium, or argon. It is pumped into the dielectric space of the line at a pressure that can vary from 3 to 35 pounds per square inch. The inert gas is used to dry the line when it is first installed and pressure is maintained to ensure that no moisture enters the line.

Flexible coaxial lines (fig. 3-14) are made with an inner conductor that consists of flexible wire insulated from the outer conductor by a solid, continuous insulating material. The outer conductor is made of metal braid, which gives the line flexibility. Early attempts at gaining flexibility involved using rubber insulators between the two conductors. However, the rubber insulators caused excessive losses at high frequencies.

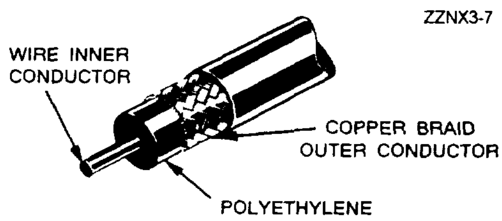

Figure 3-14.—Flexible coaxial line.

Because of the high-frequency losses associated with rubber insulators, polyethylene plastic was developed to replace rubber and eliminate these losses. Polyethylene plastic is a solid substance that remains flexible over a wide range of temperatures. It is unaffected by seawater, gasoline, oil, and most other liquids that may be found aboard ship. The use of polyethylene as an insulator results in greater high-frequency losses than the use of air as an insulator. However, these losses are still lower than the losses associated with most other solid dielectric materials.

This concludes our study of transmission lines. The rest of this chapter will be an introduction into the study of waveguides.

WAVEGUIDE THEORY

The two-wire transmission line used in conventional circuits is inefficient for transferring electromagnetic energy at microwave frequencies. At these frequencies, energy escapes by radiation because the fields are not confined in all directions, as illustrated in figure 3-15. Coaxial lines are more efficient than two-wire lines for transferring electromagnetic energy because the fields are completely confined by the conductors, as illustrated in figure 3-16. Waveguides are the most efficient way to transfer electromagnetic energy. WAVEGUIDES are essentially coaxial lines without center conductors. They are constructed from conductive material and may be rectangular, circular, or elliptical in shape, as shown in figure 3-17.

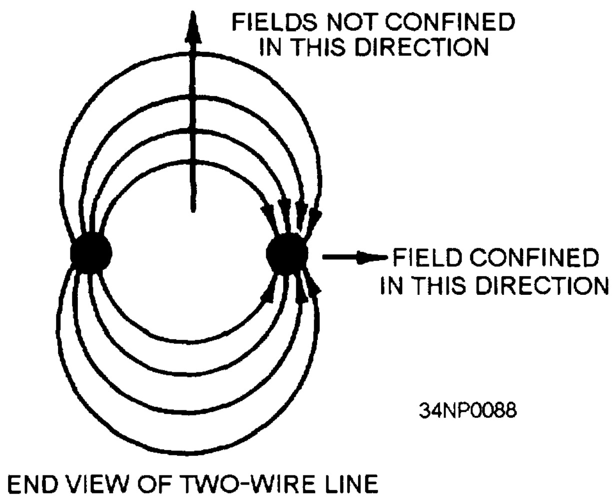

Figure 3-15.—Fields confined in two directions only.

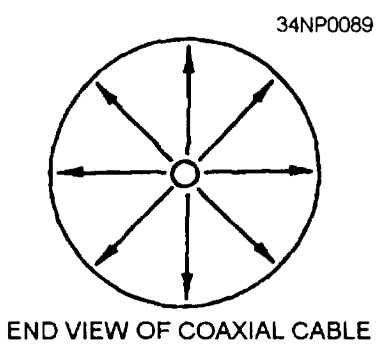

Figure 3-16.—Fields confined in all directions.

WAVEGUIDE ADVANTAGES

Waveguides have several advantages over two-wire and coaxial transmission lines. For example, the large surface area of waveguides greatly reduces COPPER (1^2R) LOSSES. Two-wire transmission lines have large copper losses because they have a relatively small surface area. The surface area of the outer conductor of a coaxial cable is large, but the surface area of the inner conductor is relatively small. At microwave frequencies, the current-carrying area of the inner conductor is restricted to a very small layer at the surface of the conductor by an action called SKIN EFFECT.

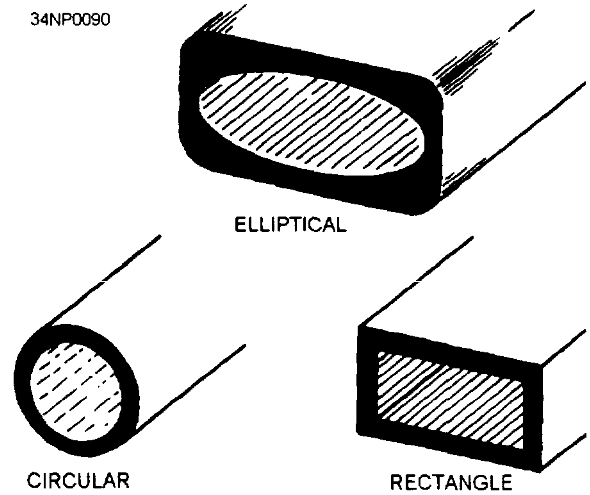

Figure 3-17.—Waveguide shapes.

Skin effect tends to increase the effective resistance of the conductor. Although energy transfer in coaxial cable is caused by electromagnetic field motion, the magnitude of the field is limited by the size of the current-carrying area of the inner conductor. The small size of the center conductor is even further reduced by skin effect, and energy transmission by coaxial cable becomes less efficient than by waveguides. DIELECTRIC LOSSES are also lower in waveguides than in two-wire and coaxial transmission lines. Dielectric losses in two-wire and coaxial lines are caused by the heating of the insulation between the conductors. The insulation behaves as the dielectric of a capacitor formed by the two wires of the transmission line. A voltage potential across the two wires causes heating of the dielectric and results in a power loss. In practical applications, the actual breakdown of the insulation between the conductors of a transmission line is more frequently a problem than is the dielectric loss.

This breakdown is usually caused by stationary voltage spikes or “nodes,” which are caused by standing waves. Standing waves are stationary and occur when part of the energy traveling down the line is reflected by an impedance mismatch with the load. The voltage potential of the standing waves at the points of greatest magnitude can become large enough to break down the insulation between transmission line conductors.

The dielectric in waveguides is air, which has a much lower dielectric loss than conventional insulating materials. However, waveguides are also subject to dielectric breakdown caused by standing waves. Standing waves in waveguides cause arcing, which decreases the efficiency of energy transfer and can severely damage the waveguide. Also since the electromagnetic fields are completely contained within the waveguide, radiation losses are kept very low.

Power-handling capability is another advantage of waveguides. Waveguides can handle more power than coaxial lines of the same size because power-handling capability is directly related to the distance between conductors. Figure 3-18 illustrates the greater distance between conductors in a waveguide.

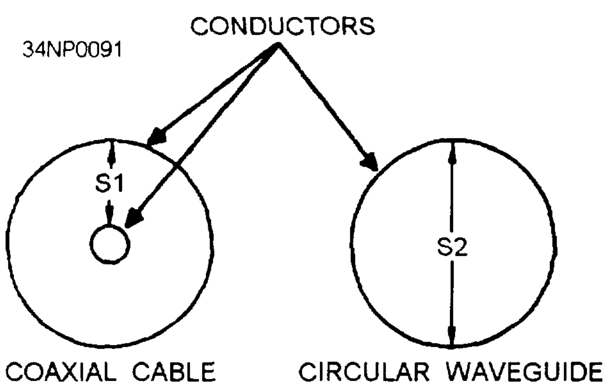

Figure 3-18.—Comparison of spacing in coaxial cable and a circular waveguide.

In view of the advantages of waveguides, you would think that waveguides should be the only type of transmission lines used. However, waveguides have certain disadvantages that make them practical for use only at microwave frequencies.

WAVEGUIDE DISADVANTAGES

Physical size is the primary lower-frequency limitation of waveguides. The width of a waveguide must be approximately a half wavelength at the frequency of the wave to be transported. For example, a waveguide for use at 1 megahertz would be about 700 feet wide. This makes the use of waveguides at frequencies below 1000 megahertz increasingly impractical. The lower frequency range of any system using waveguides is limited by the physical dimensions of the waveguides.

Waveguides are difficult to install because of their rigid, hollow-pipe shape. Special couplings at the joints are required to assure proper operation. Also, the inside surfaces of waveguides are often plated with silver or gold to reduce skin effect losses. These requirements increase the costs and decrease the practicality of waveguide systems at any other than microwave frequencies.

DEVELOPING THE WAVEGUIDE FROM PARALLEL LINES

You may better understand the transition from ordinary transmission line concepts to waveguide theories by considering the development of a waveguide from a two-wire transmission line. Figure 3-19 shows a section of a two-wire transmission line supported on two insulators. At the junction with the line, the insulators must present a very high impedance to ground for proper operation of the line. A low impedance insulator would obviously short-circuit the line to ground, and this is what happens at very high frequencies. Ordinary insulators display the characteristics of the dielectric of a capacitor formed by the wire and ground. As the frequency increases, the overall impedance decreases. A better high-frequency insulator is a quarter-wave section of transmission line shorted at one end. Such an insulator is shown in figure 3-20. The impedance of a shorted quarter-wave section is very high at the open-end junction with the two-wire transmission line. This type of insulator is known as a METALLIC INSULATOR and may be placed anywhere along a two-wire line.

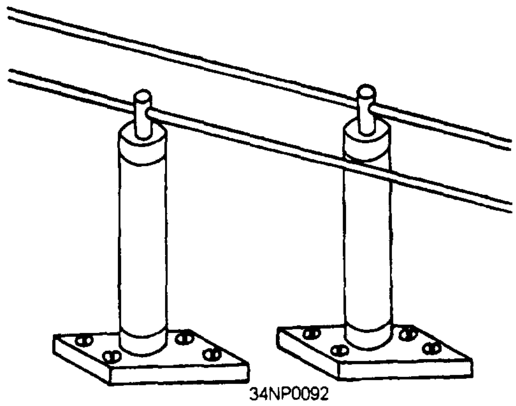

Figure 3-19.—Two-wire transmission line.

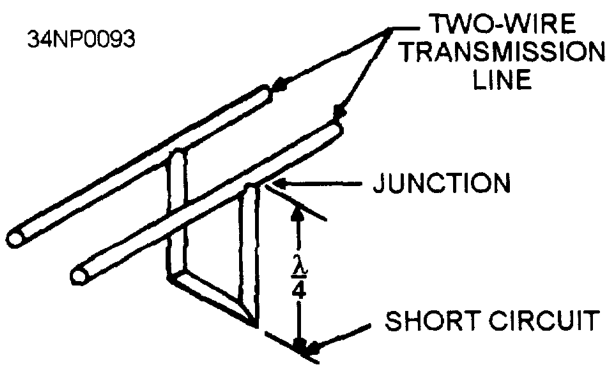

Figure 3-20.—Quarter-wave section of transmission line shorted at one end.

Note that quarter-wave sections are insulators at only one frequency. This severely limits the bandwidth, efficiency, and application of this type of two-wire line.

Figure 3-21 shows several metallic insulators on each side of a two-wire transmission line. As more insulators are added, each section makes contact with the next, and a rectangular waveguide is formed. The lines become part of the walls of the waveguide, as illustrated in figure 3-22. The energy is then conducted within the hollow waveguide instead of along the two-wire transmission line.

The comparison of the way electromagnetic fields work on a transmission line and in a waveguide is not exact. During the change from a two-wire line to a waveguide, the electromagnetic field configurations also undergo many changes. As a result of these changes, the waveguide does not actually operate

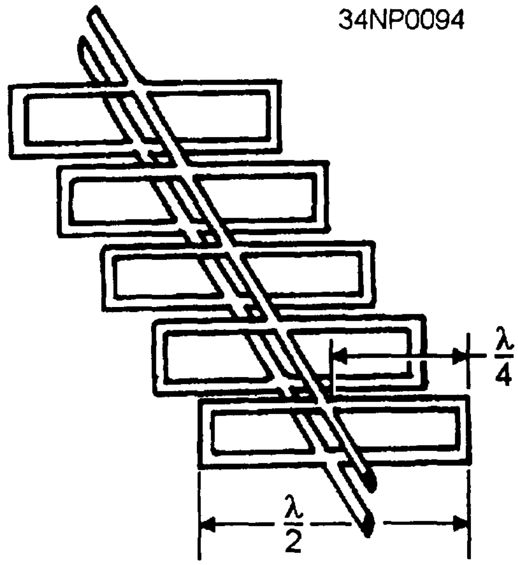

Figure 3-21.—Metallic insulator on each side of a two-wire line.

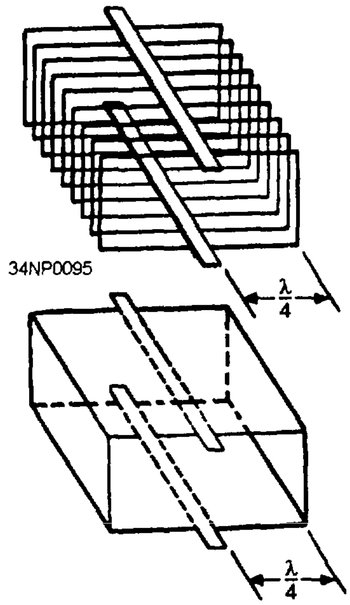

Figure 3-22.—Forming a waveguide by adding quarter-wave sections.

like a two-wire line that is completely shunted by quarter-wave sections. If it did, the use of a wave-guide would be limited to a single-frequency wave length that was four times the length of the quarter-wave sections. In fact, waves of this length cannot pass efficiently through waveguides. Only a small range of frequencies of somewhat shorter wavelength (higher frequency) can pass efficiently.

As shown in figure 3-23, the widest dimension of a waveguide is called the “a” dimension and determines the range of operating frequencies. The narrowest dimension determines the power-handling capability of the waveguide and is called the “b” dimension.

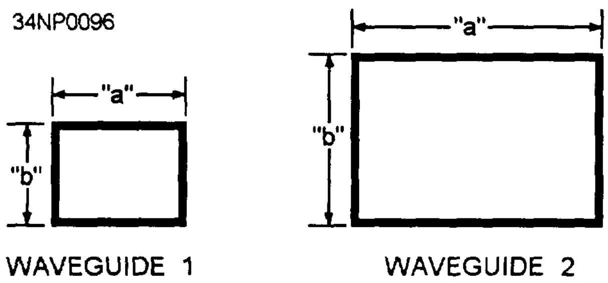

Figure 3-23.—Labeling waveguide dimensions,

NOTE: This method of labeling waveguides is not standard in all texts, Different methods may be used in other texts on microwave principles, but this method is in accordance with Navy Military Standards (MIL-STDS).

In theory, a waveguide could function at an infinite number of frequencies higher than the designed frequency; however, in practice, an upper frequency limit is caused by modes of operation, which will be discussed later.

If the frequency of a signal is decreased so much that two quarter-wavelengths are longer than the wide dimension of a waveguide, energy will no longer pass through the waveguide. This is the lower frequency limit, or CUTOFF FREQUENCY of a given waveguide. In practical applications, the wide dimension of a waveguide is usually 0.7 wavelength at the operating frequency. This allows the waveguide to handle a small range of frequencies both above and below the operating frequency. The “b” dimension is governed by the breakdown potential of the dielectric, which is usually air. Dimensions ranging from 0.2 to 0.5 wavelength are common for the “b” sides of a waveguide.

ENERGY PROPAGATION IN WAVEGUIDES

Since energy is transferred through waveguides by electromagnetic fields, you need a basic understanding of field theory. Both electric (E FIELD) and magnetic fields (H FIELD) are present in waveguides, and the interaction of these fields causes energy to travel through the waveguide. This action is best understood by first looking at the properties of the two individual fields.

E Field

An electric field exists when a difference of potential causes a stress in the dielectric between two points. The simplest electric field is one that forms between the plates of a capacitor when one plate is made positive compared to the other, as shown in view A of figure 3-24. The stress created in the dielectric is an electric field.

Electric fields are represented by arrows that point from the positive toward the negative potential. The number of arrows shows the relative strength of the field. In view B, for example, evenly spaced arrows indicate the field is evenly distributed. For ease of explanation, the electric field is abbreviated E field, and the lines of stress are called E lines.

H Field

The magnetic field in a waveguide is made up of magnetic lines of force that are caused by current flow through the conductive material of the waveguide. Magnetic lines of force, called H lines, are continuous closed loops, as shown in figure 3-25. All of the H lines associated with current are collectively called a magnetic field or H field. The strength of the H field, indicated by the number of H lines in a given area, varies directly with the amount of current.

Although H lines encircle a single, straight wire, they behave differently when the wire is formed into a coil, as shown in figure 3-26. In a coil the individual H lines tend to form around each turn of wire. Since

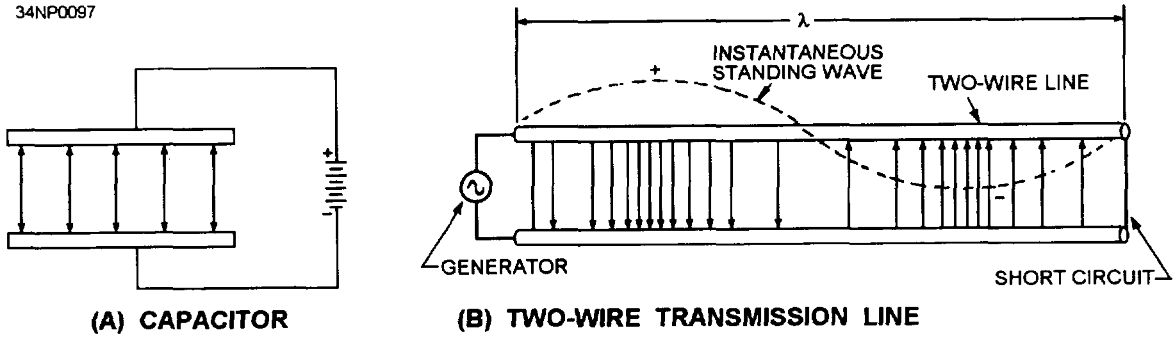

Figure 3-24.—Simple electric fields.

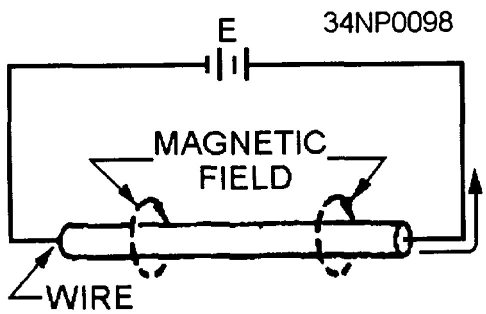

Figure 3-25.—Magnetic field on a single wire.

the H lines take opposite directions between adjacent turns, the field between the turns is canceled. Inside and outside the coil, where the direction of each H field is the same, the fields join and form continuous H lines around the entire coil. A similar action takes place in a waveguide.

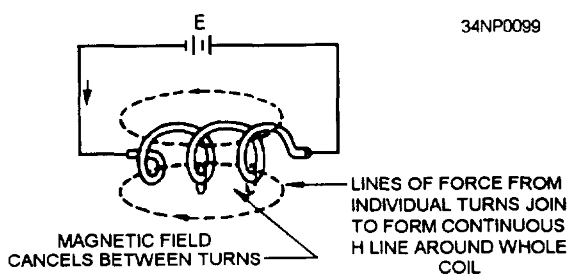

Figure 3-26.—Magnetic field on a coil.

BOUNDARY CONDITIONS IN A WAVEGUIDE

The travel of energy down a waveguide is similar, but not identical, to the travel of electromagnetic waves in free space. The difference is that the energy in a waveguide is confined to the physical limits of the guide. Two conditions, known as BOUNDARY CONDITIONS, must be satisfied for energy to travel through a waveguide.

The first boundary condition (illustrated in fig. 3-27, view A can be stated as follows:

For an electric field to exist at the surface of a conductor, it must be perpendicular to the conductor.

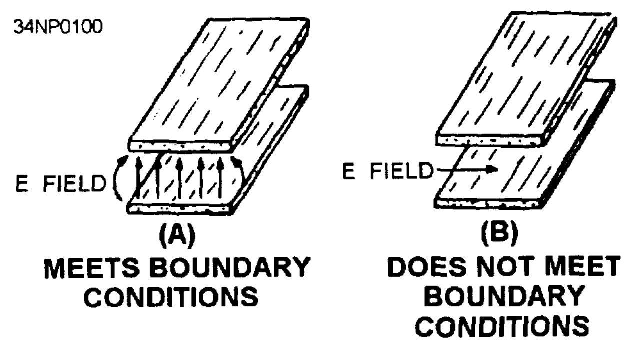

Figure 3-27.—E field boundary condition.

The opposite of this boundary condition, shown in view B, is also true. An electric field CANNOT exist parallel to a perfect conductor.

The second boundary condition, which is illustrated in figure 3-28, can be stated as follows:

For a varying magnetic field to exist, it must form closed loops in parallel with the conductors and be perpendicular to the electric field.

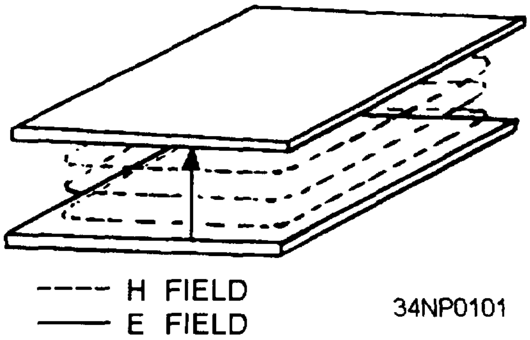

Figure 3-28.—H field boundary condition.

Since an E field causes a current flow that in turn produces an H field, both fields always exist at the same time in a waveguide. If a system satisfies one of these boundary conditions, it must also satisfy the other since neither field can exist alone.

WAVEFRONTS WITHIN A WAVEGUIDE

Electromagnetic energy transmitted into space consists of electric and magnetic fields that are at right angles (90 degrees) to each other and at right angles to the direction of propagation. A simple analogy to establish this relationship is by use of the right-hand rule for electromagnetic energy, based on the POYNTING VECTOR. It indicates that a screw (right-hand thread) with its axis perpendicular to the electric and magnetic fields will advance in the direction of propagation if the E field is rotated to the right (toward the H field). This rule is illustrated in figure 3-29.

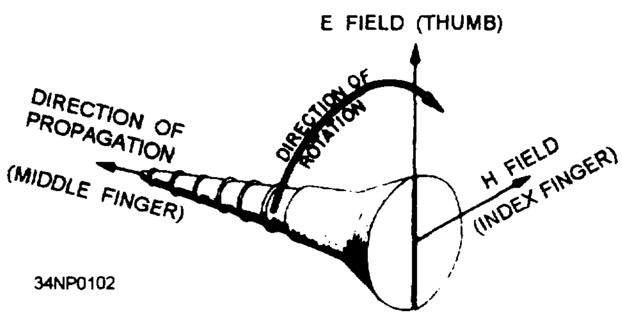

Figure 3-29.—The Poynting vector.

The combined electric and magnetic fields form a wavefront that can be represented by alternate negative and positive peaks at half-wavelength intervals, as illustrated in figure 3-30. Angle $\varnothing$ is the direction of travel of the wave with respect to some reference axis.

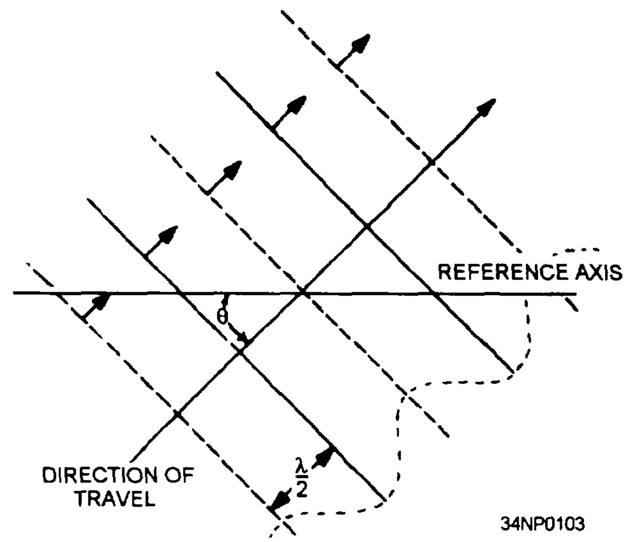

Figure 3-30.—Wavefronts in space.

The reflection of a single wavefront off the "b" wall of a waveguide is shown in figure 3-31. The wavefront is shown in view A as small particles, In views B and C particle 1 strikes the wall and is bounced back from the wall without losing velocity. If the wall is perfectly flat, the angle at which it the wall, known as the angle of incidence (θ), is the same as the angle of reflection ($\varnothing$). An instant after particle 1 strikes the wall, particle 2 strikes the wall, as shown

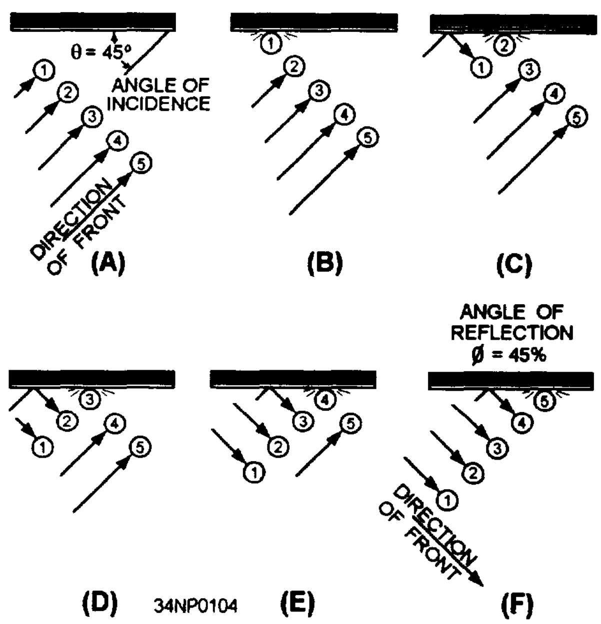

Figure 3-31.—Reflection of a single wavefront.

in view C, and reflects in the same manner. Because all the particles are traveling at the same velocity, particles 1 and 2 do not change their relative position with respect to each other. Therefore, the reflected wave has the same shape as the original. The remaining particles as shown in views D, E, and F reflect in the same manner. This process results in a reflected wavefront identical in shape, but opposite in polarity, to the incident wave.

Figure 3-32, views A and B, each illustrate the direction of propagation of two different electromagnetic wavefronts of different frequencies being radiated into a waveguide by a probe. Note that only the direction of propagation is indicated by the lines and arrowheads. The wavefronts are at right angles to the direction of propagation. The angle of incidence (θ) and the angle of reflection$(\varnothing)$ of the wavefronts vary in size with the frequency of the input energy, but the angles of reflection are equal to each other in a waveguide. The CUTOFF FREQUENCY in a waveguide is a frequency that would cause angles of incidence and reflection to be perpendicular to the walls of the guide. At any frequency below the cutoff frequency, the wavefronts will be reflected back and forth across the guide (setting up standing waves) and no energy will be conducted down the waveguide.

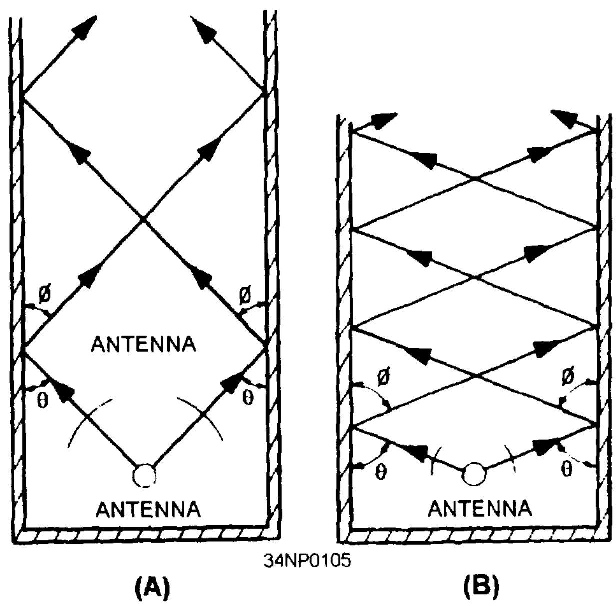

Figure 3-32.—Different frequencies in a waveguide.

The velocity of propagation of a wave along a waveguide is less than its velocity through free space (speed of light). This lower velocity is caused by the zigzag path taken by the wavefront. The forward-progress velocity of the wavefront in a waveguide is called GROUP VELOCITY and is somewhat slower than the speed of light.

The group velocity of energy in a waveguide is determined by the reflection angle of the wavefronts off the "b" walls. The reflection angle is determined by the frequency of the input energy. This basic principle is illustrated in figure 3-33. As frequency is decreased. the reflection angle increases, causing the group velocity to decrease. The opposite is also true; increasing frequency increases the group velocity.

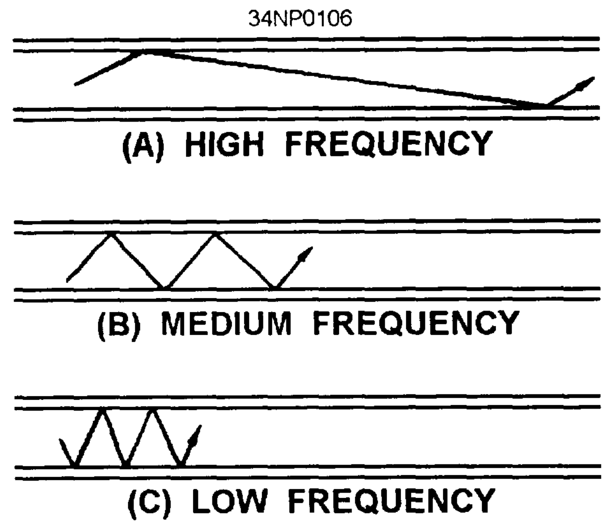

Figure 3-33.—Reflection angle at various frequencies.

WAVEGUIDE MODES OF OPERATION

The waveguide analyzed in the previous paragraphs yields an electric field configuration known as the half-sine electric distribution. This configuration, called a MODE OF OPERATION, is shown in figure 3-34. Recall that the strength of the field is indicated by the spacing of the lines; that is, the closer the lines, the stronger the field. The regions of maximum voltage in this field move continuously down the waveguide in a sine-wave pattern. To meet boundary conditions. the field must always be zero at the "b" walls.

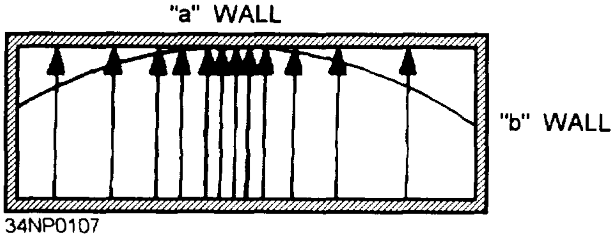

Figure 3-34.—Half-sine E field distribution.

The half-sine field is only one of many field configurations, or modes, that can exist in a rectangular waveguide. A full-sine field can also exist in a rectangular waveguide because, as shown in figure 3-35, the field is zero at the "b" walls.

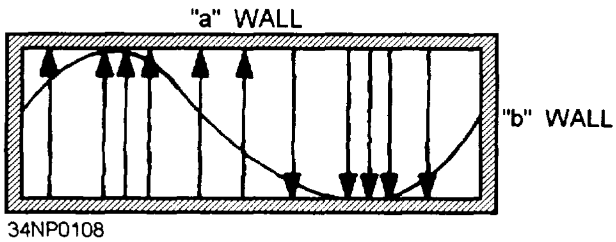

Figure 3-35.—Full-sine E field distribution.

The magnetic field in a rectangular waveguide is in the form of closed loops parallel to the surface of the conductors. The strength of the magnetic field is proportional to the electric field. Figure 3-36 illustrates the magnetic field pattern associated with a half-sine electric field distribution. The magnitude of the magnetic field varies in a sine-wave pattern down the center of the waveguide in "time phase" with the electric field. TIME PHASE means that the peak H lines and peak E lines occur at the same instant in time, although not necessarily at the same point along the length of the waveguide.

The dominant mode is the most efficient mode. Waveguides are normally designed so that only the dominant mode will be used. To operate in the dominant mode, a waveguide must have an "a" (wide) dimension of at least one half-wavelength of the frequency to be propagated. The "a" dimension of the waveguide must be kept near the minimum allowable value to ensure that only the dominant mode will exist. In practice, this dimension is usually 0.7 wavelength.

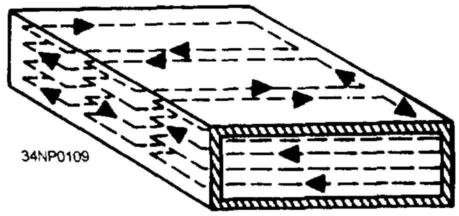

Figure 3-36.—Magnetic field caused by a half-sine E field.

Of the possible modes of operation available for a given waveguide, the dominant mode has the lowest cutoff frequency. The high-frequency limit of a rectangular waveguide is a frequency at which its "a" dimension becomes large enough to allow operation in a mode higher than that for which the waveguide has been designed.

Circular waveguides are used in specific areas of radar and communications systems, such as rotating joints used at the mechanical point where the antennas rotate. Figure 3-37 illustrates the dominant mode of a circular waveguide. The cutoff wavelength of a circular guide is 1.71 times the diameter of the waveguide. Since the "a" dimension of a rectangular waveguide is approximately one half-wavelength at the cutoff frequency, the diameter of an equivalent circular waveguide must be 2/1.71, or approximately

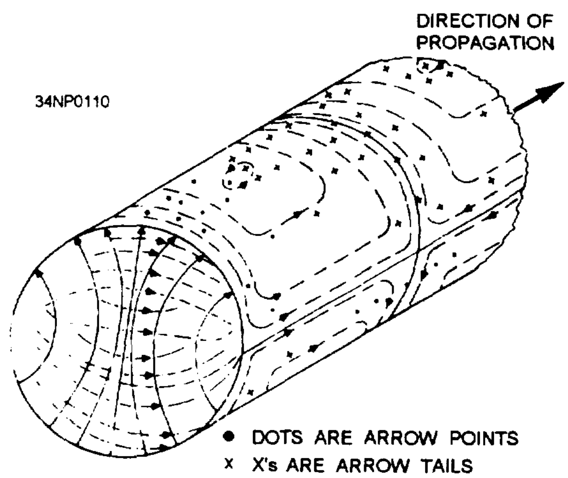

Figure 3-37.—Dominant mode in a circular waveguide.

1.17 times the "a" dimension of a rectangular waveguide.

MODE NUMBERING SYSTEMS

So far, only the most basic types of E and H field arrangements have been shown. More complicated arrangements are often necessary to make possible coupling, isolation, or other types of operation. The field arrangements of the various modes of operation are divided into two categories: TRANSVERSE ELECTRIC (TE) and TRANSVERSE MAGNETIC (TM).

In the transverse electric (TE) mode, the entire electric field is in the transverse plane, which is perpendicular to the waveguide, (direction of energy travel). Part of the magnetic field is parallel to the length axis.

In the transverse magnetic (TM) mode, the entire magnetic field is in the transverse plane and has no portion parallel to the length axis.

Since there are several TE and TM modes, subscripts are used to complete the description of the field pattern. In rectangular waveguides, the first subscript indicates the number of half-wave patterns in the "a" dimension, and the second subscript indicates the number of half-wave patterns in the "b" dimension.

The dominant mode for rectangular waveguides is shown in figure 3-38. It is designated as the TE mode because the E fields are perpendicular to the "a" walls. The first subscript is 1, since there is only one half-wave pattern across the "a" dimension. There are no E-field patterns across the "b" dimension, so the second subscript is 0. The complete mode description of the dominant mode in rectangular waveguides is $TE_{1,0}$. Subsequent description of waveguide operation in this text will assume the dominant ($TE_{1,0}$) mode unless otherwise noted.

A similar system is used to identify the modes of circular waveguides. The general classification of TE and TM is true for both circular and rectangular waveguides. In circular waveguides the subscripts have a different meaning. The first subscript indicates the number of fill-wave patterns around the circumference of the waveguide. The second subscript indicates the number of half-wave patterns across the diameter.

In the circular waveguide in figure 3-39, the E field is perpendicular to the length of the waveguide with no E lines parallel to the direction of propagation. Thus, it must be classified as operating in the TE mode. If you follow the E line pattern in a counter-clockwise direction starting at the top, the E lines go from zero, through maximum positive (tail of arrows), back to zero, through maximum negative (head of arrows), and then back to zero again. This is one full wave, so the first subscript is 1. Along the diameter, the E lines go from zero through maximum and back to zero, making a half-wave variation. The second subscript, therefore, is also 1. $TE_{1,1}$ is the complete mode description of the dominant mode in circular waveguides. Several modes are possible in both circular and rectangular waveguides. Figure 3-40 illustrates several different modes that can be used to verify the mode numbering system.

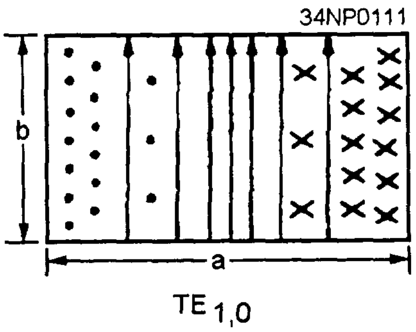

Figure 3-38.—Dominant mode in a rectangular waveguide.

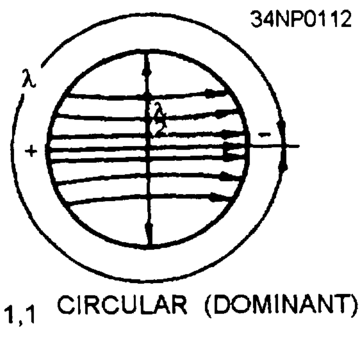

Figure 3-39.—Counting wavelengths in a circular waveguide.

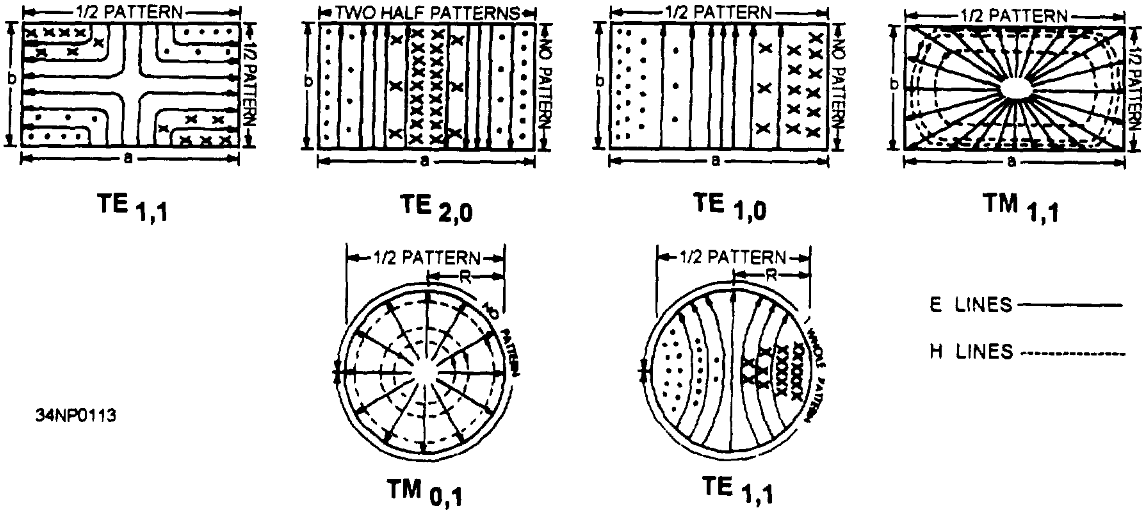

Figure 3-40.—Various modes of operation for rectangular and circular waveguides.

WAVEGUIDE INPUT/OUTPUT METHODS

A waveguide, as explained earlier in this topic, operates differently from an ordinary transmission line. Therefore, special devices must be used to put energy into a waveguide at one end and remove it from the other end.

The three devices used to injector remove energy from waveguides are PROBES, LOOPS, and SLOTS. Slots may also be called APERTURES or WINDOWS.

When a small probe is inserted into a waveguide and supplied with microwave energy, it acts as a quarter-wave antenna. Current flows in the probe and sets up an E field such as the one shown in figure 3-41, view A. The E lines detach themselves from the probe. When the probe is located at the point of highest efficiency, the E lines set up an E field of considerable intensity.

The most efficient place to locate the probe is in the center of the “a” wall, parallel to the “b” wall, and one quarter-wavelength from the shorted end of the waveguide, as shown in figure 3-41, views B and C. This is the point at which the E field is maximum in the dominant mode. Therefore, energy transfer (coupling) is maximum at this point. Note that the quarter-wavelength spacing is at the frequency required to propagate the dominant mode.

In many applications a lesser degree of energy transfer, called loose coupling, is desirable. The amount of energy transfer can be reduced by decreasing the length of the probe, by moving it out of the center of the E field, or by shielding it. Where the degree of coupling must be varied frequently, the probe is made retractable so the length can be easily changed.

The size and shape of the probe determines its frequency, bandwidth, and power-handling capability. As the diameter of a probe increases, the bandwidth increases. A probe similar in shape to a door knob is capable of handling much higher power and a larger bandwidth than a conventional probe. The greater power-handling capability is directly related to the increased surface area. Two examples of broad-bandwidth probes are illustrated in figure 3-41, view D. Removal of energy from a waveguide is simply a reversal of the injection process using the same type of probe.

Another way of injecting energy into a waveguide is by setting up an H field in the waveguide. This can be accomplished by inserting a small loop that carries a high current into the waveguide, as shown in figure 3-42, view A. A magnetic field builds up around the loop and expands to fit the waveguide, as shown in view B. If the frequency of the current in the loop is within the bandwidth of the waveguide, energy will be transferred to the waveguide.

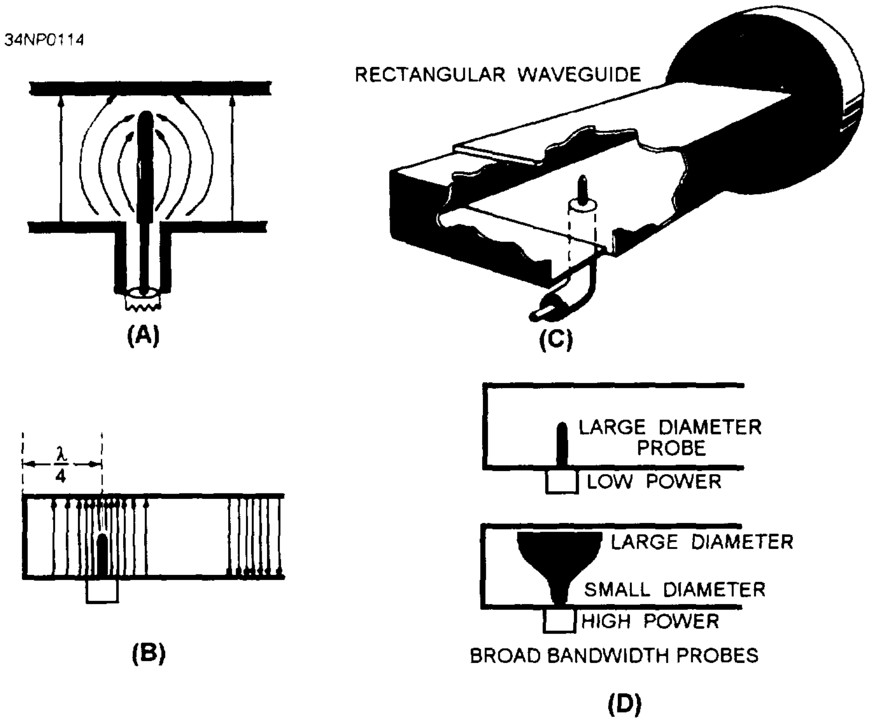

Figure 3-41.—Probe coupling in a rectangular waveguide.

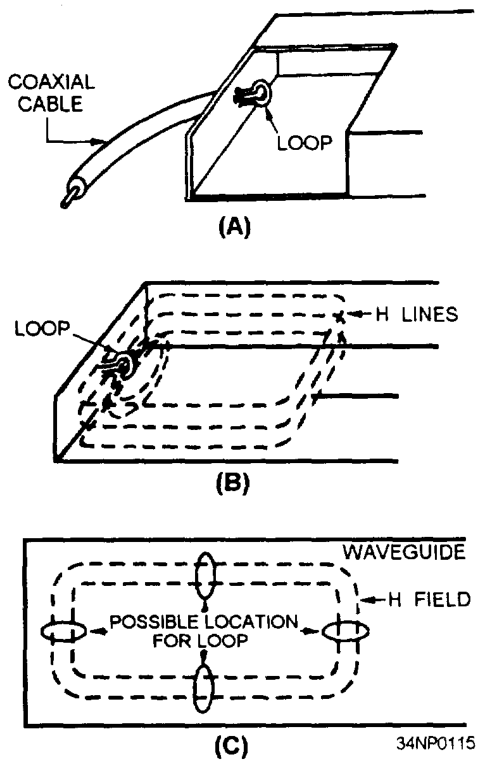

Figure 3-42.—Loop coupling in a rectangular waveguide.

For the most efficient coupling to the waveguide, the loop is inserted at one of several points where the magnetic field will be of greatest strength. Four of those points are shown in figure 3-42, view C.

When less efficient coupling is desired, you can rotate or move the loop until it encircles a smaller number of H lines. When the diameter of the loop is increased, its power-handling capability also increases. The bandwidth can be increased by increasing the size of the wire used to make the loop.

When a loop is introduced into a waveguide in which an H field is present, a current is induced in the loop. When this condition exists, energy is removed from the waveguide.

Slots or apertures are sometimes used when very loose (inefficient) coupling is desired, as shown in figure 3-43. In this method energy enters through a small slot in the waveguide and the E field expands into the waveguide. The E lines expand first across the slot and then across the interior of the waveguide.

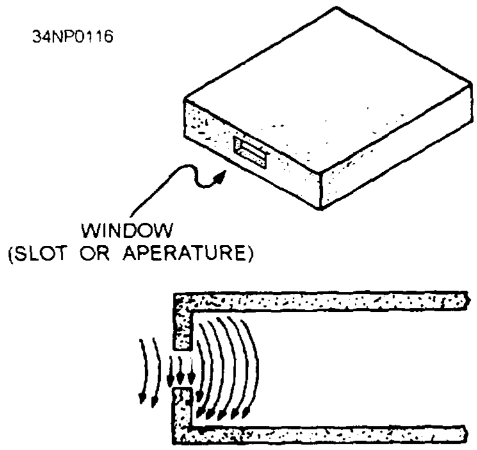

Figure 3-43.—Slot coupling in a waveguide.

Minimum reflections occur when energy is injected or removed if the size of the slot is properly proportioned to the frequency of the energy.

After learning how energy is coupled into and out of a waveguide with slots, you might think that leaving the end open is the most simple way of injecting or removing energy in a waveguide. This is not the case, however, because when energy leaves a waveguide, fields form around the end of the waveguide. These fields cause an impedance mismatch which, in turn, causes the development of standing waves and a drastic loss in efficiency. Various methods of impedance matching and terminating waveguides will be covered in the next section.

WAVEGUIDE IMPEDANCE MATCHING

Waveguide transmission systems are not always perfectly impedance matched to their load devices. The standing waves that result from a mismatch cause a power loss, a reduction in power-handling capability, and an increase in frequency sensitivity. Impedance-changing devices are therefore placed in the waveguide to match the waveguide to the load. These devices are placed near the source of the standing waves.

Figure 3-44 illustrates three devices, called irises, that are used to introduce inductance or capacitance into a waveguide. An iris is nothing more than a metal plate that contains an opening through which the waves may pass. The iris is located in the transverse plane of either the magnetic or electric field.

An inductive iris and its equivalent circuit are illustrated in figure 3-44, view A. The iris places a shunt inductive reactance across the waveguide that is directly proportional to the size of the opening. Notice that the inductive iris is in the magnetic plane. The shunt capacitive reactance, illustrated in view B, basically acts the same way. Again, the reactance is directly proportional to the size of the opening, but the iris is placed in the electric plane. The iris, illustrated in view C, has portions in both the magnetic

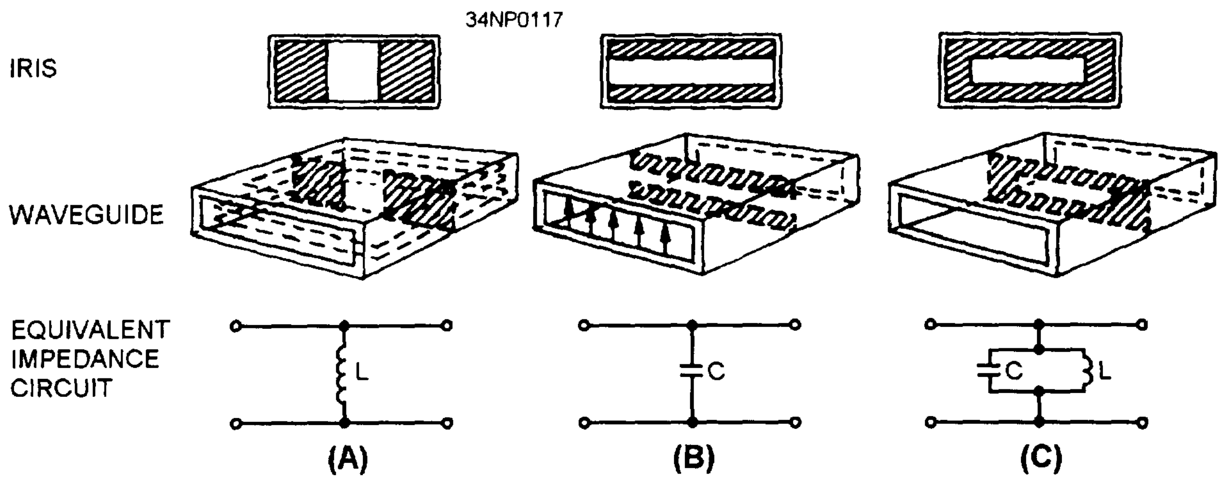

Figure 3-44.—Waveguide irises.

and electric transverse planes and forms an equivalent parallel-LC circuit across the waveguide. At the resonant frequency, the iris acts as a high shunt resistance. Above or below resonance, the iris acts as a capacitive or inductive reactance.

POSTS and SCREWS made from conductive material can be used for impedance-changing devices in waveguides. Views A and B of figure 3-45, illustrate two basic methods of using posts and screws. A post or screw that only partially penetrates into the waveguide acts as a shunt capacitive reactance. When the post or screw extends completely through the waveguide, making contact with the top and bottom walls, it acts as an inductive reactance. Note that when screws are used, the amount of reactance can be varied.

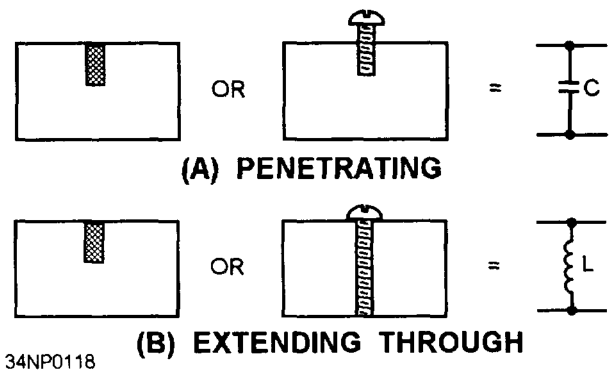

Figure 3-45.—Conducting posts and screws.

WAVEGUIDE TERMINATIONS

Electromagnetic energy is often passed through a waveguide to transfer the energy from a source into space. As previously mentioned, the impedance of a waveguide does not match the impedance of space, and without proper impedance matching standing waves cause a large decrease in the efficiency of the waveguide.

Any abrupt change in impedance causes standing waves, but when the change in impedance at the end of a waveguide is gradual, almost no standing waves are formed. Gradual changes in impedance can be obtained by terminating the waveguide with a funnel-shaped HORN, such as the three types illustrated in figure 3-46. The type of horn used depends upon the frequency and the desired radiation pattern.

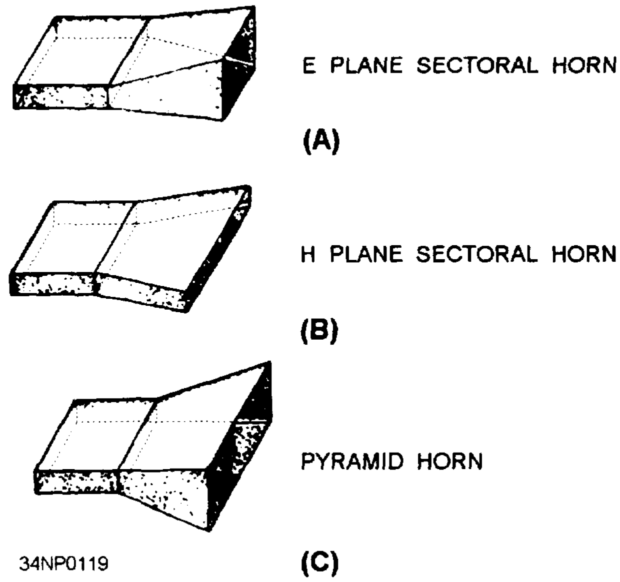

Figure 3-46.—Waveguide horns.

As you may have noticed, horns are really simple antennas. They have several advantages over other impedance-matching devices, such as their large bandwidth and simple construction.

A waveguide may also be terminated in a resistive load that is matched to the characteristic impedance of the waveguide. The resistive load is most often called a DUMMY LOAD, because its only purpose is to absorb all the energy in a waveguide without causing standing waves.

There is no place on a waveguide to connect a fixed termination resistor; therefore, several special arrangements are used to terminate waveguides. One method is to fill the end of the waveguide with a graphite and sand mixture, as illustrated in figure 3-47, view A. When the fields enter the mixture, they induce a current flow in the mixture that dissipates the energy as heat. Another method (view B) is to use a high-resistance rod placed at the center of the E field. The E field causes current to flow in the rod, and the high resistance of the rod dissipates the energy as a power loss, again in the form of heat.

Still another method for terminating a waveguide is the use of a wedge of highly resistive material, as shown in view C of figure 3-47. The plane of the wedge is placed perpendicular to the magnetic lines

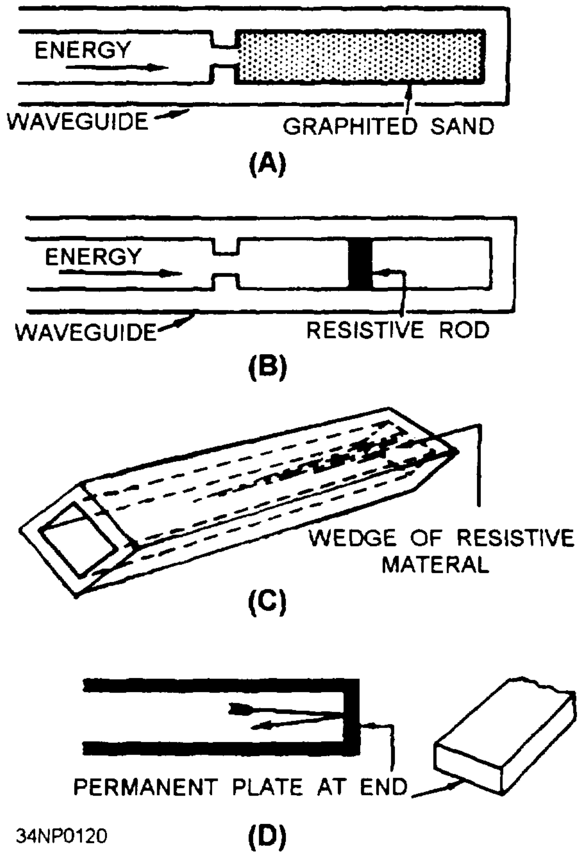

Figure 3-47.—Terminating waveguides.

of force. When the H lines cut through the wedge, current flows in the wedge and causes a power loss. As with the other methods, this loss is in the form of heat. Since very little energy reaches the end of the waveguide, reflections are minimum.

All of the terminations discussed so far are designed to radiate or absorb the energy without reflections. In many instances, however, all of the energy must be reflected from the end of the waveguide. The best way to accomplish this is to permanently weld a metal plate at the end of the waveguide, as shown in view D of figure 3-47.

WAVEGUIDE PLUMBING

Since waveguides are really only hollow metal pipes, the installation and the physical handling of waveguides have many similarities to ordinary plumbing. In light of this fact, the bending, twisting, joining, and installation of waveguides is commonly called waveguide plumbing. Naturally, waveguides are different in design from pipes that are designed to carry liquids or other substances. The design of a waveguide is determined by the frequency and power level of the electromagnetic energy it will carry. The following paragraphs explain the physical factors involved in the design of waveguides.

Waveguide Bends

The size, shape, and dielectric material of a waveguide must be constant throughout its length for energy to move from one end to the other without reflections. Any abrupt change in its size or shape can cause reflections and a loss in overall efficiency. When such a change is necessary, the bends, twists, and joints of the waveguides must meet certain conditions to prevent reflections.

Waveguides maybe bent in several ways that do not cause reflections. One way is the gradual bend shown in figure 3-48. This gradual bend is known as an E bend because it distorts the E fields. The E bend must have a radius greater than two wavelengths to prevent reflections.

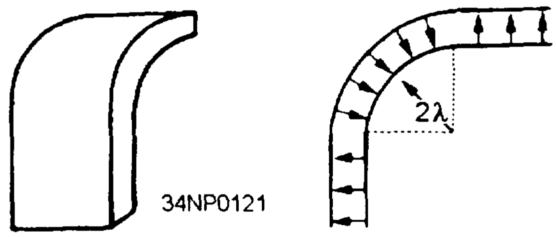

Figure 3-48.—Gradual E bend.

Another common bend is the gradual H bend (fig. 3-49). It is called an H bend because the H fields are distorted when a waveguide is bent in this manner. Again, the radius of the bend must be greater than two wavelengths to prevent reflections. Neither the E bend in the “a” dimension nor the H bend in the “b” dimension changes the normal mode of operation.

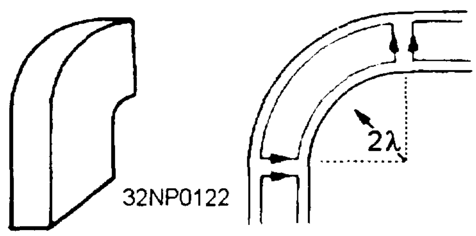

Figure 3-49.—Gradual H bend.

A sharp bend in either dimension may be used if it meets certain requirements. Notice the two 45-degree bends in figure 3-50; the bends are $1/4\lambda$ apart. The reflections that occur at the 45-degree bends cancel each other, leaving the fields as though no reflections have occurred.

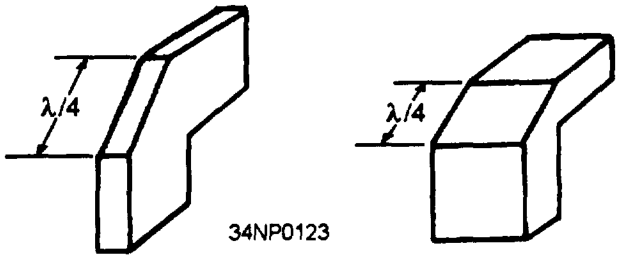

Figure 3-50.—Sharp bends.

Sometimes the electromagnetic fields must be rotated so that they are in the proper phase to match the phase of the load. This may be accomplished by twisting the waveguide as shown in figure 3-51. The twist must be gradual and greater than 2λ.

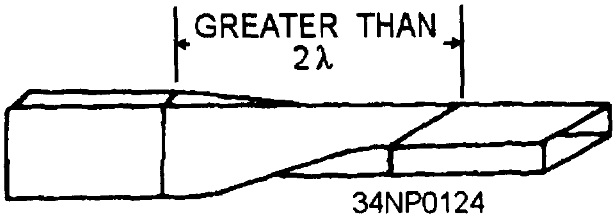

Figure 3-51.—Waveguide twist.

The flexible waveguide (fig. 3-52) allows special bends, which some equipment applications might require. It consists of a specially wound ribbon of conductive material, the most commonly used is brass, with the inner surface plated with chromium. Power losses are greater in the flexible waveguide because the inner surfaces are not perfectly smooth. Therefore, it is only used in short sections where no other reasonable solution is available.

Waveguide Joints

Since an entire waveguide system cannot possibly be molded into one piece, the waveguide must be

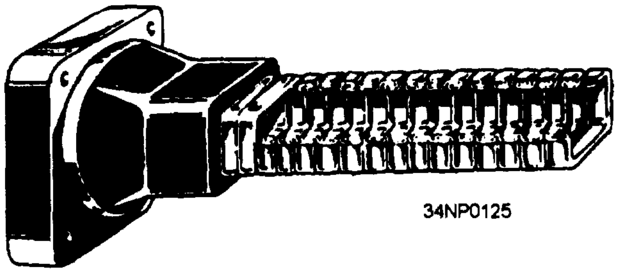

Figure 3-52.—Flexible waveguide.

constructed in sections and the sections connected with joints. The three basic types of waveguide joints are the PERMANENT, the SEMIPERMANENT, and the ROTATING JOINTS. Since the permanent joint is a factory-welded joint that requires no maintenance, only the semipermanent and rotating joints will be discussed.

Sections of waveguide must be taken apart for maintenance and repair. A semipermanent joint, called a CHOKE JOINT, is most commonly used for this purpose. The choke joint provides good electromagnetic continuity between the sections of the waveguide with very little power loss.

A cross-sectional view of a choke joint is shown in figure 3-53. The pressure gasket shown between the two metal surfaces forms an airtight seal. Notice in view B that the slot is exactly $1/4\lambda$ from the "a" wall of the waveguide. The slot is also $1/4\lambda$ deep, as shown in view A, and because it is shorted at point 1, a high impedance results at point 2. Point 3 is $1/4\lambda$ from point 2. The high impedance at point 2 results in a low impedance, or short, at point 3. This effect creates a good electrical connection between the two sections that permits energy to pass with very little reflection or loss.

Whenever a stationary rectangular waveguide is to be connected to a rotating antenna, a rotating joint must be used. A circular waveguide is normally used in a rotating joint. Rotating a rectangular waveguide would cause field pattern distortion. The rotating section of the joint, illustrated in figure 3-54, uses a choke joint to complete the electrical connection with the stationary section. The circular waveguide is designed so that it will operate in the $TM_{0,1}$ mode.

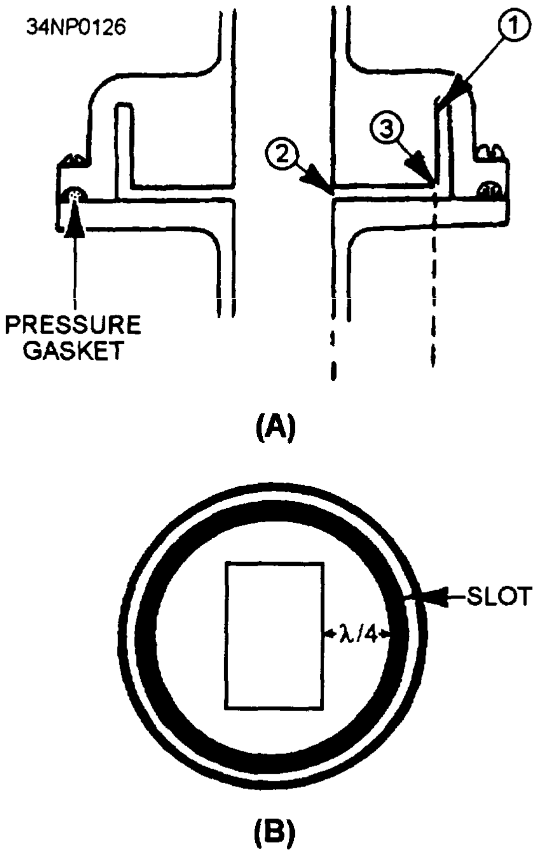

Figure 3-53.—Choke joint.

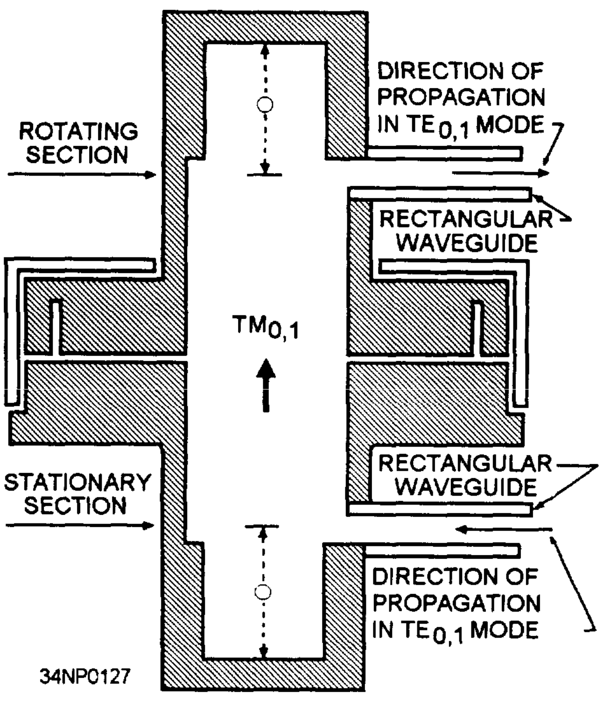

Figure 3-54.—Rotating joint.

The rectangular sections are attached as shown in the illustration to prevent the circular waveguide from operating in the wrong mode. Distance "O" is $3/4\lambda$ so that a high impedance will be presented to any unwanted modes. This is the most common design used for rotating joints, but other types may be used in specific applications.

WAVEGUIDE MAINTENANCE

The installation of a waveguide system presents problems that are not normally encountered when dealing with other types of transmission lines. These problems often fall within the technician's area of responsibility. A brief discussion of waveguide handling, installation, and maintenance will help prepare you for this maintenance responsibility, Detailed information concerning waveguide maintenance in a particular system may be found in the technical manuals for the system.

Since a waveguide naturally has a low loss ratio, most losses in a waveguide system are caused by other factors. Improperly connected joints or damaged inner surfaces can decrease the efficiency of a system to the point that it will not work at all. Therefore, you must take great care when working with waveguides to prevent physical damage. Since waveguides are made from a soft, conductive material, such as copper or aluminum, they are very easy to dent or deform. Even the slightest damage to the inner surface of a waveguide will cause standing waves and, often, internal arcing. Internal arcing causes further damage to the waveguide in an action that is often self-sustaining until the waveguide is damaged beyond use. Part of your job as a technician will be to inspect the waveguide system for physical damage. The previously mentioned dents are only one type of physical damage that can decrease the efficiency of the system. Another problem occurs because waveguides are made from a conductive material such as copper while the structures of most ships are made from steel. When two dissimilar metals, such as copper and steel, are in direct contact, an electrical action called ELECTROLYSIS takes place that causes very rapid corrosion of the metals. Waveguides can be completely destroyed by electrolytic corrosion in a relatively short period of time if they are not isolated from direct contact with other metals. Any inspection of a waveguide system should include a detailed inspection of all support points to ensure that electrolytic corrosion is not taking place. Any waveguide that is exposed to the weather should be painted and all joints sealed. Proper painting prevents natural corrosion, and sealing the joints prevents moisture from entering the waveguide.

Moisture can be one of the worst enemies of a waveguide system. As previously discussed, the dielectric in waveguides is air, which is an excellent dielectric as long as it is free of moisture. Wet air, however, is a very poor dielectric and can cause serious internal arcing in a waveguide system. For this reason, care is taken to ensure that waveguide systems are pressurized with air that is dry. Checking the pressure and moisture content of the waveguide air may be one of your daily system maintenance duties.

More detailed waveguide installation and maintenance information can be found in the technical manuals that apply to your particular system. Another good source is the *Electronics Installation and Maintenance Handbooks (EIMB)* published by Naval Sea Systems Command. *Installation Standards (EIMB) Handbook,* NAVSEA 0967-LP-000-0110, is the volume that deals with waveguide installation and maintenance.

WAVEGUIDE DEVICES

The discussion of waveguides, up to this point, has been concerned only with the transfer of energy from one point to another. Many waveguide devices have been developed, however, that modify the energy in some fashion during the transmission. Some devices do nothing more than change the direction of the energy. Others have been designed to change the basic characteristics or power level of the electromagnetic energy.

This section will explain the basic operating principles of some of the more common waveguide devices, such as DIRECTIONAL COUPLERS, CAVITY RESONATORS, and HYBRID JUNCTIONS.

Directional Couplers

The directional coupler is a device that provides a method of sampling energy from within a waveguide for measurement or use in another circuit. Most couplers sample energy traveling in one direction only. However, directional couplers can be constructed that sample energy in both directions. These are called BIDIRECTIONAL couplers and are widely used in radar and communications systems.

Directional couplers may be constructed in many ways. The coupler illustrated in figure 3-55 is constructed from an enclosed waveguide section of the same dimensions as the waveguide in which the energy is to be sampled. The “b” wall of this enclosed section is mounted to the “b” wall of the waveguide from which the sample will be taken. There are two holes in the “b” wall between the sections of the coupler. These two holes are $1/4\lambda$ apart. The upper section of the directional coupler has a wedge of energy-absorbing material at one end and a pickup probe connected to an output jack at the other end. The absorbent material absorbs the energy not directed at the probe and a portion of the overall energy that enters the section.

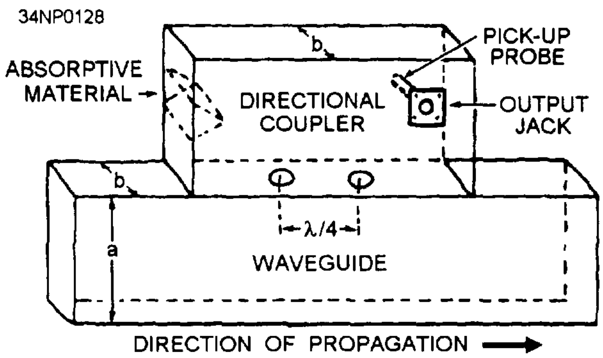

Figure 3-55.—Directional coupler.

Figure 3-56 illustrates two portions of the incident wavefront in a waveguide. The waves travel down the waveguide in the direction indicated and enter the coupler section through both holes. Since both portions of the wave travel the same distance, they are in phase when they arrive at the pickup probe. Because the waves are in phase, they add together and provide a sample of the energy traveling down the waveguide. The sample taken is only a small portion of the energy that is traveling down the waveguide. The magnitude of the sample, however, is proportional to the magnitude of the energy in the waveguide. The absorbent material is designed to ensure that the ratio between the sample energy and the energy in the waveguide is constant. Otherwise, the sample would contain no useful information. The ratio is usually stamped on the coupler in the form of an attenuation factor.

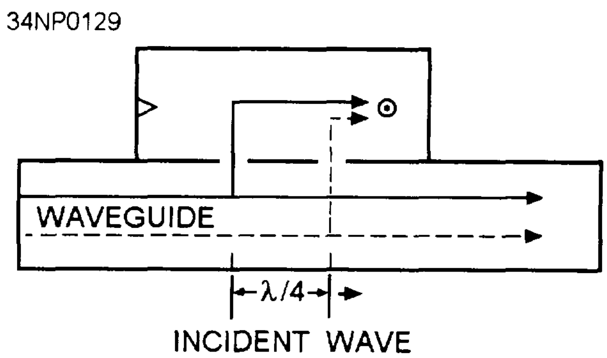

Figure 3-56.—Incident wave in a directional coupler designed to sample incident waves.

The effect of a directional coupler on any reflected energy is illustrated in figure 3-57. Note that these two waves do not travel the same distance to the pickup probe. The wave represented by the dotted line travels $1/2\lambda$ further and arrives at the probe 180 degrees out of phase with the wave, represented by the solid line. Because the waves are 180 degrees out of phase at the probe, they cancel each other and no energy is induced into the pickup probe. When the reflected energy arrives at the absorbent material, it adds and is absorbed by the material.

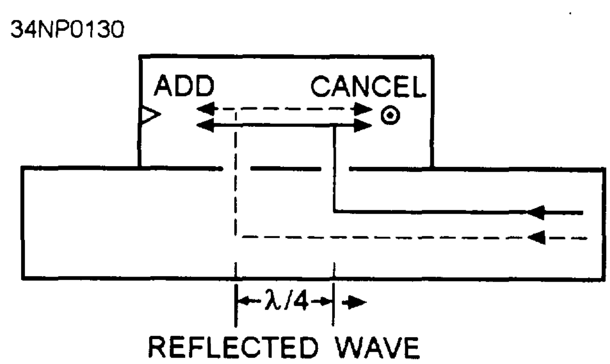

Figure 3-57.—Reflected wave in a directional coupler.

A directional coupler designed to sample reflected energy is shown in figure 3-58. The absorbent material and the probe are in opposite positions from the directional coupler designed to sample the incident energy. This positioning causes the two portions of the reflected energy to arrive at the probe in phase, providing a sample of the reflected energy. The transmitted energy is absorbed by the absorbent material.

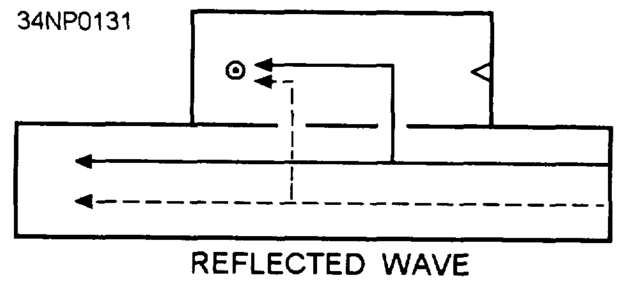

Figure 3-58.—Directional coupler designed to sample retlected energy.

A simple bidirectional coupler for sampling both transmitted and reflected energy can be constructed by mounting two directional couplers on opposite sides of a waveguide, as shown in figure 3-59.

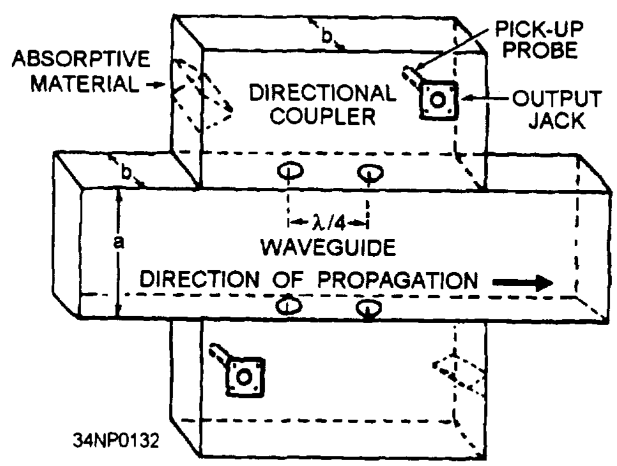

Figure 3-59.—Bidirectional coupler.

Cavity Resonators

By definition, a resonant cavity is any space completely enclosed by conducting walls that can contain oscillating electromagnetic fields and possess resonant properties. The cavity has many advantages and uses at microwave frequencies. Resonant cavities have a very high Q and can be built to handle relatively large amounts of power. Cavities with a Q value in excess of 30,000 are not uncommon. The high Q gives these devices a narrow bandpass and allows very accurate tuning. Simple, rugged construction is an additional advantage.

Although cavity resonators, built for different frequency ranges and applications, have a variety of shapes, the basic principles of operation are the same for all.

One example of a cavity resonator is the rectangular box shown in figure 3-60, view A. It may be thought of as a section of rectangular waveguide closed at both ends by conducting plates. The frequency at which the resonant mode occurs is $1/2\lambda$ of the distance between the end plates. The magnetic field patterns in the rectangular cavity are shown in view B.

There are two variables that determine the primary frequency of any resonant cavity. The first variable is PHYSICAL SIZE. In general, the smaller the cavity, the higher its resonant frequency. The second controlling factor is the SHAPE of the cavity. Figure 3-61 illustrates several cavity shapes that are commonly used. Remember from the previously stated definition of a resonant cavity that any completely enclosed conductive surface, regardless of its shape, can act as a cavity resonator.

Energy can be inserted or removed from a cavity by the same methods that are used to couple energy into and out of waveguides. The operating principles of probes, loops, and slots are the same whether used in a cavity or a waveguide. Therefore, any of the three methods can be used with cavities to inject or remove energy.

The resonant frequency of a cavity can be varied by changing any of the three parameters: cavity volume, cavity capacitance, or cavity inductance. Changing the frequencies of a cavity is known as TUNING. The mechanical methods of tuning a cavity may vary with the application, but all methods use the same electrical principles.

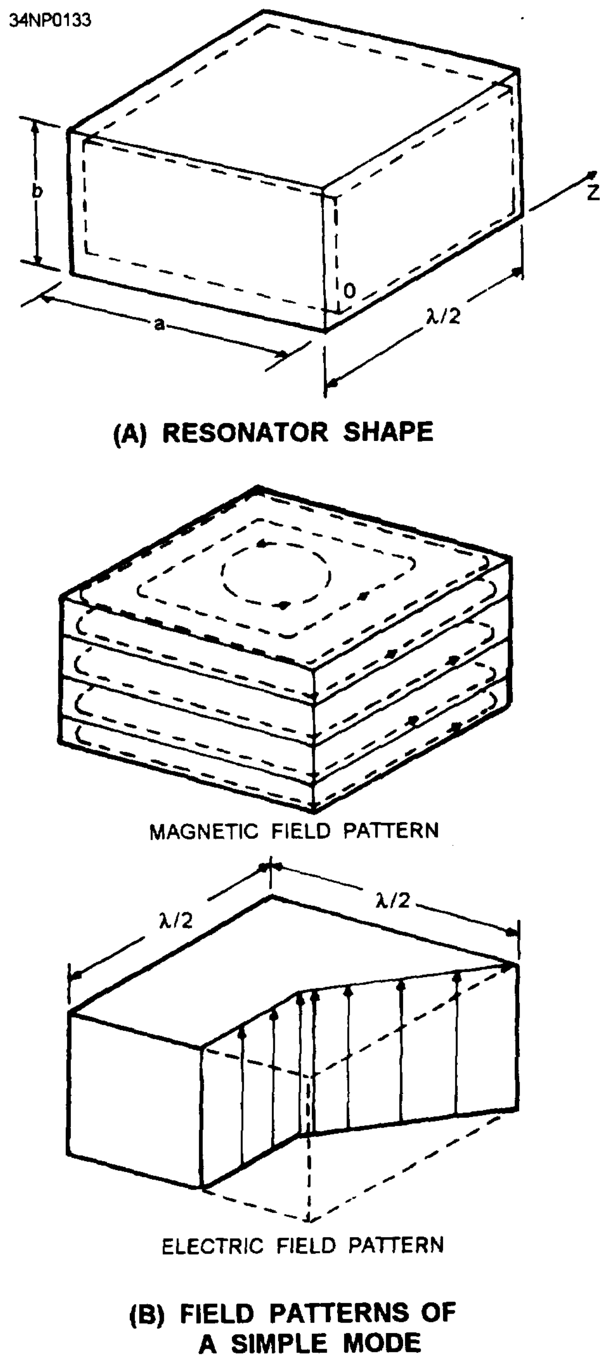

Figure 3-60.—Rectangular waveguide cavity resonator.

Waveguide Junctions

You may have assumed that when energy traveling down a waveguide reaches a junction it simply divides and follows the junction. This is not strictly true.

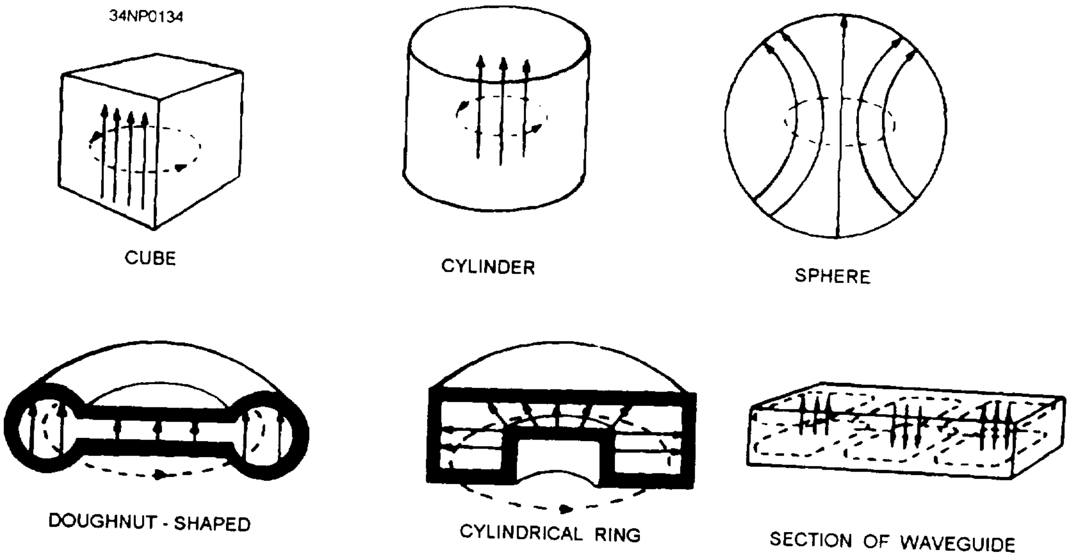

Figure 3-61.—Types of cavities.

Different types of junctions affect the energy in different ways. Since waveguide junctions are used extensively in most systems, you need to understand the basic operating principles of those most commonly used.

The T JUNCTION is the most simple of the commonly used waveguide junctions. T junctions are divided into two basic types, the E TYPE and the H TYPE. HYBRID JUNCTIONS are more complicated developments of the basic T junctions. The MAGIC-T and the HYBRID RING are the two most commonly used hybrid junctions.

E-TYPE T JUNCTION.— An E-type T junction is illustrated in figure 3-62, view A.

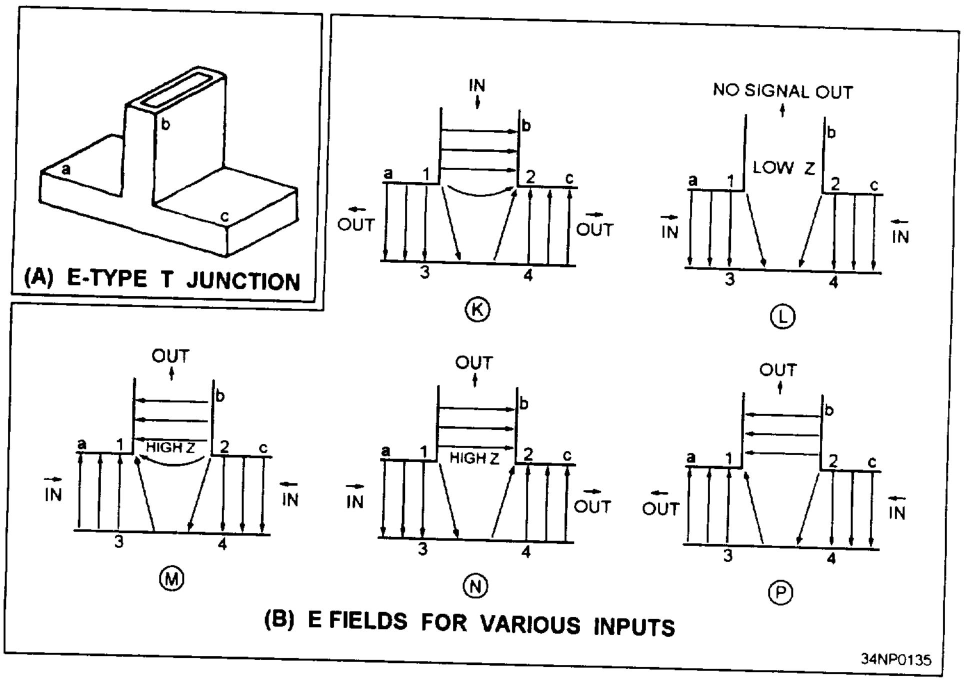

Figure 3-62.—E fields in an E-type T junction.

It is called an E-type T junction because the junction arm extends from the main waveguide in the same direction as the E field in the waveguide.

Figure 3-62, view B, illustrates cross-sectional views of the E-type T junction with inputs fed into the various arms. For simplicity, the magnetic lines that are always present with an electric field have been omitted. In view K, the input is fed into arm b and the outputs are taken from the a and c arms. When the E field arrives between points 1 and 2, point 1 becomes positive and point 2 becomes negative. The positive charge at point 1 then induces a negative charge on the wall at point 3. The negative charge at point 2 induces a positive charge at point 4. These charges cause the fields to form 180 degrees out of phase in the main waveguide; therefore, the outputs will be 180 degrees out of phase with each other. In view L, two in-phase inputs of equal amplitude are fed into the a and c arms. The signals at points 1 and 2 have the same phase and amplitude. No difference of potential exists across the entrance to the b arm, and no energy will be coupled out. However, when the two signals fed into the a and c arms are 180 degrees out of phase, as shown in view M, points 1 and 2 have a difference of potential. This difference of potential induces an E field from point 1 to point 2 in the b arm, and energy is coupled out of this arm. Views N and P illustrate two methods of obtaining two outputs with only one input.

H-TYPE T JUNCTION.— An H-type T junction is illustrated in figure 3-63, view A. It is called an H-type T junction because the long axis of the "b" arm is parallel to the plane of the magnetic lines of force in the waveguide. Again, for simplicity, only the E lines are shown in this figure. Each X indicates an E line moving away from the observer. Each dot indicates an E line moving toward the observer.

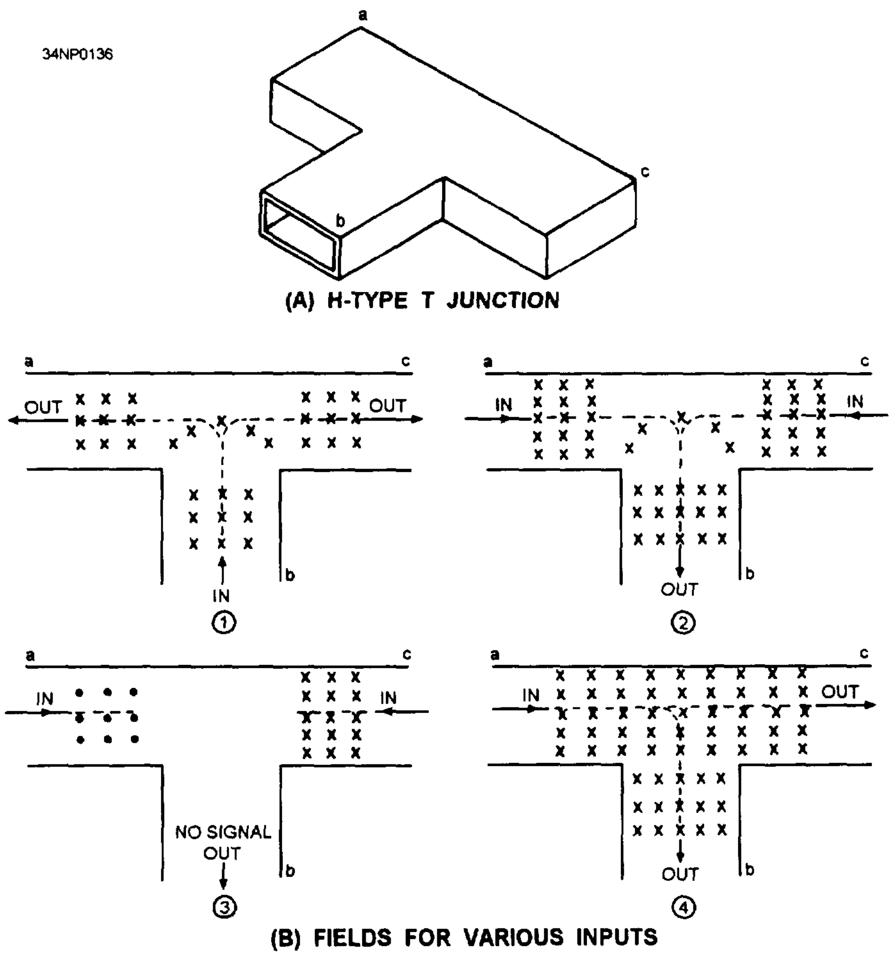

Figure 3-63.—E field in an H-type T junction.

In view 1 of figure 3-63, view B, the signal is fed into arm b and in-phase outputs are obtained from the a and c arms. In view 2, in-phase signals are fed into arms a and c and the output signal is obtained from the b arm because the fields add at the junction and induce E lines into the b arm. If 180-degree-out-of-phase signals are fed into arms a and c, as shown in view 3, no output is obtained from the b arm because the opposing fields cancel at the junction. If a signal is fed into the a arm, as shown in view 4 , outputs will be obtained from the b and c arms. The reverse is also true. If a signal is fed into the c arm, outputs will be obtained from the a and b arms.

MAGIC-T HYBRID JUNCTION.— A simplified version of the magic-T hybrid junction is shown in figure 3-64. The magic-T is a combination of the H-type and E-type T junctions. The most common application of this type of junction is as the mixer section for microwave radar receivers.

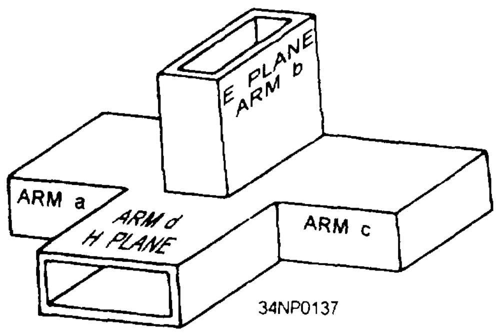

Figure 3-64.—Magic-T hybrid junction.

If a signal is fed into the b arm of the magic-T, it will divide into two out-of-phase components. As shown in figure 3-65, view A, these two components will move into the a and c arms. The signal entering the b arm will not enter the d arm because of the zero potential existing at the entrance of the d arm. The potential must be zero at this point to satisfy the boundary conditions of the b arm. This absence of potential is illustrated in views B and C where the magnitude of the E field in the b arm is indicated by the length of the arrows. Since the E lines are at maximum in the center of the b arm and minimum at the edge where the d arm entrance is located, no potential difference exists across the mouth of the d arm.

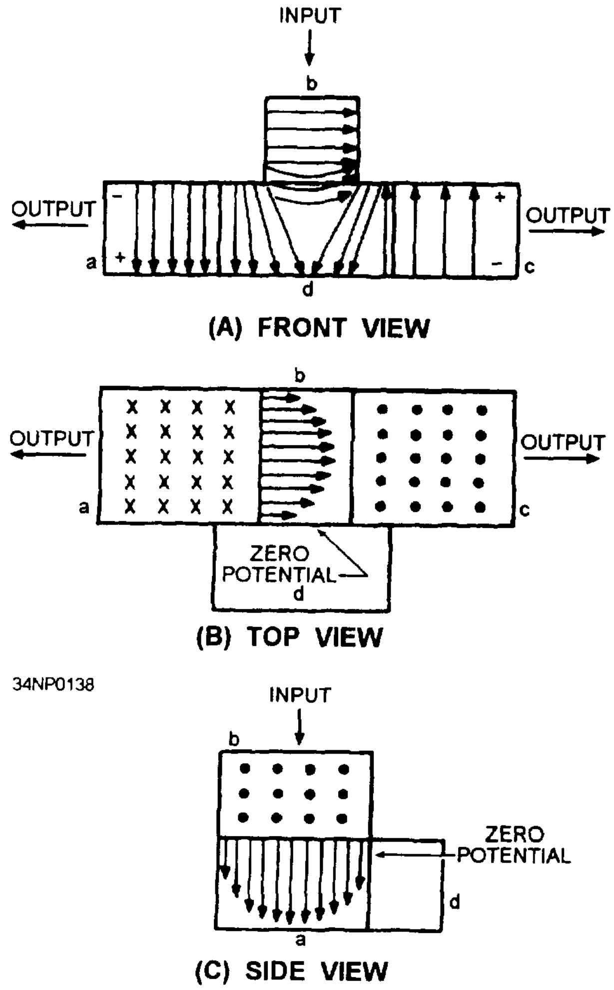

Figure 3-65.—Magic-T with input to arm b.

In summary, when an input is applied to arm b of the magic-T hybrid junction, the output signals from arms a and c are 180 degrees out of phase with each other, and no output occurs at the d arm.

The action that occurs when a signal is fed into the d arm of the magic-T is illustrated in figure 3-66. As with the H-type T junction, the signal entering the d arm divides and moves down the a and c arms as outputs that are in phase with each other and with the input. The shape of the E fields in motion is shown by the numbered curved slices. As the E field moves down the d arm, points 2 and 3 are at an equal potential. The energy divides equally into arms a and c, and the E fields in both arms become identical in shape. Since the potentials on both sides of the b arm are equal, no potential difference exists at the entrance to the b arm, resulting in no output.

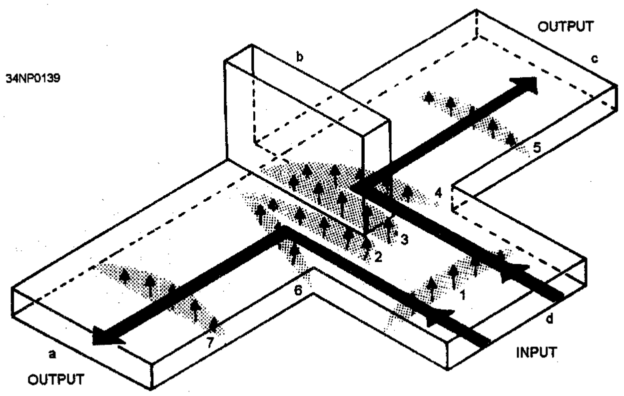

Figure 3-66.—Magic-T with input to arm d.

When an input signal is fed into the a arm as shown in figure 3-67, a portion of the energy is coupled into the b arm as it would be in an E-type T junction. An equal portion of the signal is coupled through the d arm because of the action of the H-type junction. The c arm has two fields across it that are out of phase with each other. Therefore, the fields cancel, resulting in no output at the c arm. The reverse of this action takes place if a signal is fed into the c arm, resulting in outputs at the b and d arms and no output at the a arm.

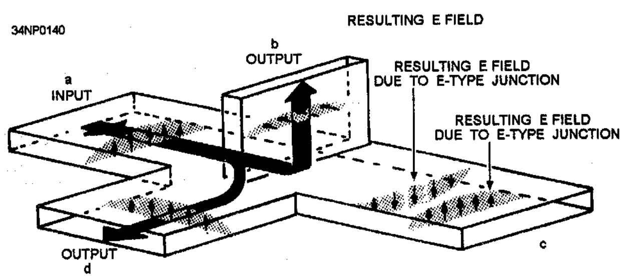

Figure 3-67.—Magic-T with input to arm a.

Unfortunately, when a signal is applied to any arm of a magic-T, the flow of energy in the output arms is affected by reflections. Reflections are caused by impedance mismatching at the junctions. These reflections are the cause of the two major disadvantages of the magic-T. First, the reflections represent a power loss since all the energy fed into the junction does not reach the load that the arms feed. Second, the reflections produce standing waves that can result in internal arcing. Thus, the maximum power a magic-T can handle is greatly reduced.

Reflections can be reduced by using some means of impedance matching that does not destroy the shape of the junctions. One method is shown in figure 3-68. A post is used to match the H plane, and an iris is used to match the E plane. Even though this method reduces reflections, it lowers the power-handling capability even further.

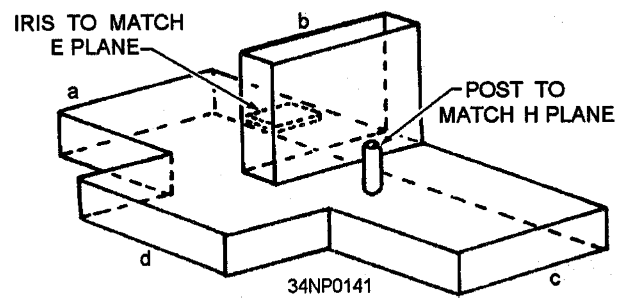

Figure 3-68.—Magic-T impedance matching.

HYBRID RING.— A type of hybrid junction that overcomes the power limitation of the magic-T is the hybrid ring, also called a RAT RACE. The hybrid ring, illustrated in figure 3-69, view A, is actually a modification of the magic-T. It is constructed of rectangular waveguides molded into a circular pattern. The arms are joined to the circular waveguide to form E-type T junctions. View B shows, in wavelengths, the dimensions required for a hybrid ring to operate properly.

The hybrid ring is used primarily in high-powered radar and communications systems to perform two functions. During the transmit period, the hybrid ring couples microwave energy from the transmitter to the antenna and allows no energy to reach the receiver. During the receive cycle, the hybrid ring couples energy from the antenna to the receiver and allows no energy to reach the transmitter. Any device that performs both of these functions is called a DUPLEXER. A duplexer permits a system to use the same antenna for both transmitting and receiving.

SUMMARY

This concludes our discussion on transmission lines and waveguides. In this volume you have been given a basic introduction on wave propagation from the time it leaves the transmitter to the point of reception. In volume 8 you will be introduced to a variety of electronic support systems.

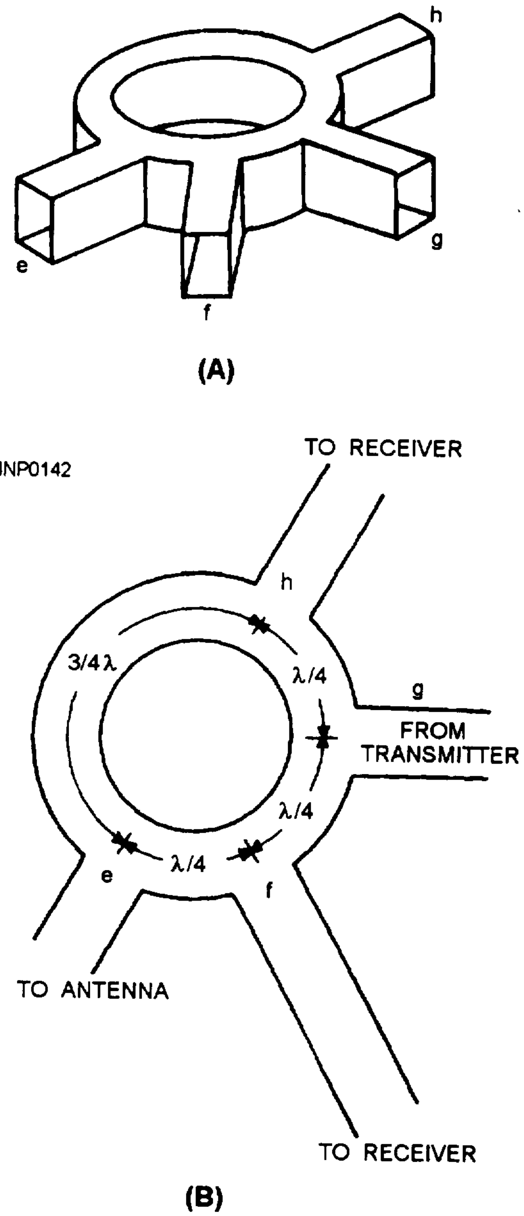

Figure 3-69.—Hybrid ring with wavelength measurements.